

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2003, 10:39:30 ; Search time 1.38425 Seconds
(without alignments)
599.261 Million cell updates/sec

Title: US-09-917-372-1_COPY_216_235
Perfect score: 108
Sequence: 1 PLPPPMGSLKRRPQGECP 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80	74.1	435	1 TNR3_HUMAN	P36941 homo sapien
2	51	47.2	897	1 AAC2_CHICK	P20111 gallus gall
3	49	45.4	172	1 VP19_CRV	P17457 cymbidium r
4	49	45.4	2142	1 BAT2_HUMAN	P48634 homo sapien
5	48	44.4	568	1 RGS3_MOUSE	Q9dc04 mus musculu
6	47	43.5	147	1 YAU8_SCHPO	Q10163 schizosacch
7	47	43.5	340	1 HEM6_SYNY3	P72848 synecocyst
8	46	42.6	173	1 VP19_CNV	P15184 cucumber ne
9	46	42.6	1124	1 JAK3_HUMAN	P52333 homo sapien
10	45	41.7	415	1 TNR3_MOUSE	P50284 mus musculu
11	44	40.7	477	1 DXR_ARATH	Q9xf59 arabidopsis
12	44	40.7	502	1 RM02_ORYSA	P92812 oryza sativ
13	44	40.7	584	1 VATA_METHH	O27036 methanobact
14	44	40.7	1132	1 BAT3_HUMAN	P46379 homo sapien
15	43.5	40.3	1733	1 VNDA_PPVKA	P33485 pseudorabie
16	43.5	40.3	2364	1 PGCA_BOVIN	P13608 bos taurus
17	43	39.8	165	1 RS10_MOUSE	P09900 mus musculu
18	43	39.8	451	1 MB31_ARATH	O04309 arabidopsis
19	43	39.8	788	1 CY14_NEUCR	P23622 neurospora
20	43	39.8	894	1 AAC2_HUMAN	P35609 homo sapien
21	43	39.8	894	1 AAC2_MOUSE	Q9ji91 mus musculu
22	43	39.8	1213	1 FMN_CHICK	Q05858 gallus gall
23	42	38.9	166	1 REL1_PANTR	P51454 pan troglod
24	42	38.9	185	1 REL1_HUMAN	P04808 homo sapien
25	42	38.9	193	1 TNR7_HUMAN	P32370 homo sapien
26	42	38.9	208	1 YS12_SYNP7	P33772 synecococc
27	42	38.9	418	1 CD15_HUMAN	Q14004 homo sapien
28	42	38.9	436	1 RHO_AQUAE	O67031 aquifex aeo
29	42	38.9	656	1 DNAA_STRCO	P27902 streptomyce
30	42	38.9	733	1 NIBL_HUMAN	Q96tal homo sapien
31	42	38.9	752	1 CO2_HUMAN	P06681 homo sapien
32	42	38.9	826	1 RGS5_HUMAN	O43374 homo sapien
33	42	38.9	894	1 MTP_HUMAN	P55157 homo sapien

34	42	38.9	1075	1 CNRA_ALCEU	P37372 alcaligenes
35	42	38.9	1076	1 NCCA_ALCXX	Q44586 alcaligenes
36	42	38.9	1100	1 JAK3_RAT	Q63272 rattus norv
37	42	38.9	1299	1 JAK3_MOUSE	Q62137 mus musculu
38	42	38.9	2774	1 MAPA_RAT	P34266 rattus norv
39	41.5	38.4	487	1 Y442_MYCTU	Q9wvs8 mus musculu
40	41.5	38.4	806	1 MK07_MOUSE	Q13164 homo sapien
41	41.5	38.4	815	1 MK07_HUMAN	Q04637 homo sapien
42	41.5	38.4	1395	1 IF4G_HUMAN	Q28343 canis famil
43	41.5	38.4	2333	1 PGCA_CANFA	O77302 lumbricus r
44	41	38.0	156	1 RS10_LUMRU	P14353 human spuma
45	41	38.0	205	1 BEL1_FOAMY	

ALIGNMENTS

RESULT 1
TNR3_HUMAN
ID TNR3_HUMAN STANDARD; PRT; 435 AA.
AC P36941:
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor receptor superfamily member 3 precursor
DE (lymphotoxin-beta receptor) (Tumor necrosis factor receptor 2 related
DE protein) (Tumor necrosis factor C receptor).
GN LTBR OR TNFRSF3 OR TNFR.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93252381; PubMed=8486360;
RA Baens M., Chaffanet M., Cassiman J.J., den Berghe H., Marynen P.;
RT "Construction and evaluation of a hncDNA library of human 12p
RT transcribed sequences derived from a somatic cell hybrid.;
RL Genomics 16:214-218(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX Strausberg R.;
RT Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP FUNCTION.
RX MEDLINE=94225209; PubMed=8171323;
RA Crowe P.D., van Arsdale T.L., Walter B.N., Ware C.F., Hession C.,
RA Ehrenfels B., Browning J.L., Din W.S., Goodwin R.G., Smith C.A.;
RT "A lymphotoxin-beta-specific receptor.;
RL Science 264:707-710(1994).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=99223511; PubMed=10207006;
RA Wu M.-Y., Wang P.-Y., Han S.-H., Hsieh S.-L.;
RT "The cytoplasmic domain of the lymphotoxin-beta receptor mediates cell
RT death in HeLa cells.;
RL J. Biol. Chem. 274:11868-11873(1999).
RN [5]
RP FUNCTION.
RX MEDLINE=20261554; PubMed=10799510;
RA Rooney I.A., Butrovich K.D., Glass A.A., Borboroglu S., Benedict C.A.,
RA Whitbeck J.C., Cohen G.H., Eisenberg R.J., Ware C.F.;
RT "The lymphotoxin-beta receptor is necessary and sufficient for
RT LIGHT-mediated apoptosis of tumor cells.;
RL J. Biol. Chem. 275:14307-14315(2000).
CC -!- FUNCTION: Receptor for the heterotrimeric lymphotoxin containing
CC LTA and LTβ, and for TNFSF14/LIGHT. Promotes apoptosis via TRAF3
CC and TRAF5. May play a role in the development of lymphoid organs.
CC -!- SUBUNIT: Self-associates.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DR EMBL; L04270; AAA36757.1; -;
 DR EMBL; BC026262; AAH26262.1; -;
 DR HSSP; P25942; ICDP.
 DR Genew; HGNC:6718; LTBR.
 DR MIM; 600979; -;
 DR InterPro; IPR001368; TNFR_c6.
 DR Pfam; PF00020; TNFR_c6; 4.
 DR ProDom; PD000771; TNFR_c6; 1.
 DR SMART; SM00208; TNFR; 4.
 DR PROSITE; PS00652; TNFR_NGFR_1; 2.
 DR PROSITE; PS00050; TNFR_NGFR_2; 3.
 KW Receptor; Apoptosis; Transmembrane; Glycoprotein; Repeat; Signal.
 FT SIGNAL 1 30 POTENTIAL.
 FT CHAIN 31 435 TUMOR NECROSIS FACTOR RECEPTOR
 FT DOMAIN 31 227 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 228 248 SUPERFAMILY MEMBER 3.
 FT DOMAIN 249 435 POTENTIAL.
 FT REPEAT 42 81 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 82 124 TNFR-CYS 1.
 FT REPEAT 125 168 TNFR-CYS 2.
 FT REPEAT 169 211 TNFR-CYS 3.
 FT REPEAT 212 255 TNFR-CYS 4.
 FT DISULFID 43 58 BY SIMILARITY.
 FT DISULFID 59 72 BY SIMILARITY.
 FT DISULFID 62 80 BY SIMILARITY.
 FT DISULFID 83 98 BY SIMILARITY.
 FT DISULFID 101 116 BY SIMILARITY.
 FT DISULFID 104 124 BY SIMILARITY.
 FT DISULFID 126 132 BY SIMILARITY.
 FT DISULFID 139 148 BY SIMILARITY.
 FT DISULFID 142 167 BY SIMILARITY.
 FT DISULFID 170 185 BY SIMILARITY.
 FT CARBOHYD 40 40 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 177 177 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 435 AA; 46709 MW; 624626E6022F656F CRC64;

Query Match 74.1%; Score 80; DB 1; Length 435;
 Best Local Similarity 35.7%; Pred. No. 0.00019;
 Matches 20; Conservative 0; Mismatches 0; Indels 36; Gaps 1;

OY 1 PLPPEMS-----GSLKRRPQGE 20
 DB 216 PLPPEMSGTMLMLAVLLPLAFLLLATVFCIWKSHPSLCRLKLSLLKRRPQGE 271
 |||||
 |||||

RESULT 2
 AAC2_CHICK STANDARD; PRT; 897 AA.
 AC P20111;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Alpha-actinin 2 (Alpha actinin skeletal muscle isoform 2) (F-actin
 cross linking protein).
 GN ACTN2.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89064821; PubMed=3197725;
 RA Arimura C., Suzuki T., Yanagisawa M., Imamura M., Hamada Y.,

Masaki T.;
 "Primary structure of chicken skeletal muscle and fibroblast alpha-
 actinins deduced from cDNA sequences";
 Eur. J. Biochem. 177:649-655(1988).
 [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=92070385; PubMed=1720388;
 RX Tokue Y., Goto S., Imamura M., Obinata T., Masaki T., Endo T.;
 "Transfection of chicken skeletal muscle alpha-actinin cDNA into
 nonmuscle and myogenic cells: dimerization is not essential for
 alpha-actinin to bind to microfilaments";
 Exp. Cell Res. 197:158-167(1991).
 RL Exp. Cell Res. 197:158-167(1991).
 CC -!- FUNCTION: F-ACTIN CROSS-LINKING PROTEIN WHICH IS THOUGHT TO ANCHOR
 ACTIN TO A VARIETY OF INTRACELLULAR STRUCTURES. THIS IS A BUNDLING
 PROTEIN.
 CC -!- SUBUNIT: HOMODIMER, ANTIPARALLEL.
 CC -!- SIMILARITY: CONTAINS 1 ACTIN-BINDING DOMAIN.
 CC -!- SIMILARITY: CONTAINS 2 CALPONIN-HOMOLOGY (CH) DOMAINS.
 CC -!- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
 CC -!- SIMILARITY: CONTAINS 4 SPECTRIN REPEATS.
 CC -----
 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DR EMBL; X13874; CAA32078.1; -;
 DR EMBL; X59247; CAA41935.1; -;
 DR PIR; S02032; S02032.
 DR PIR; S15481; S15481.
 DR HSSP; O01082; 1BKR.
 DR InterPro; IPR001589; Actbind_actnin.
 DR InterPro; IPR001715; Calponin-like.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR002017; Spectrin.
 DR Pfam; PF00036; efhand; 2.
 DR Pfam; PF00307; CH; 2.
 DR Pfam; PF00435; spectrin; 4.
 DR ProDom; PD000012; EF-hand; 1.
 DR SMART; SM00033; CH; 2.
 DR SMART; SM00054; EFh; 2.
 DR SMART; SM00150; SPEC; 2.
 DR PROSITE; PS00019; ACTININ_1; 1.
 DR PROSITE; PS00020; ACTININ_2; 1.
 DR PROSITE; PS00021; CH; 2.
 DR PROSITE; PS00018; EF_HAND; FALSE_NEG.
 KW Actin-binding; Calcium-binding; Repeat; Multigene family.
 FT DOMAIN 1 257 ACTIN-BINDING.
 FT DOMAIN 41 145 CH 1.
 FT DOMAIN 154 257 CH 2.
 FT REPEAT 284 394 SPECTRIN 1.
 FT REPEAT 404 509 SPECTRIN 2.
 FT REPEAT 519 630 SPECTRIN 3.
 FT REPEAT 640 743 SPECTRIN 4.
 FT CA_BIND 769 780 EF-HAND 1 (POTENTIAL).
 FT CA_BIND 805 816 EF-HAND 2 (POTENTIAL).
 SQ SEQUENCE 897 AA; 104275 MW; F4FAC12E7F4C8634 CRC64;
 Query Match 47.2%; Score 51; DB 1; Length 897;
 Best Local Similarity 52.6%; Pred. No. 7.6;
 Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 OY 2 LPPEMSGSLKRRPQGE 20
 DB 857 LPPEQAQYCIKRMPOYTG 875
 ||||| : |||||
 ||||| : |||||
 RESULT 3
 RP19_CRV STANDARD; PRT; 172 AA.
 ID VP19_CRV

[illegible]

KW Hypothetical protein; Nuclear protein; Ribonucleoprotein; RNA-binding.
SQ SEQUENCE 147 AA; 15476 MW; 09FD4F2D583F0FF5 CRC64;

Query Match 43.5%; Score 47; DB 1; Length 147;
Best Local Similarity 47.8%; Pred. No. 4.3;
Matches 11; Conservative 1; Mismatches 7; Indels 4; Gaps 1;

OY 1 PLPPEMSGSLK----RRPQEG 19
I: I I I I I I I I I I
Db 85 PMPDSMRGSLSGPGVAPAG 107

RESULT 7

HEM6_SYNY3
ID HEM6_SYNY3 STANDARD; PRT; 340 AA.
AC P72848;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Coproporphyrinogen III oxidase, aerobic (EC 1.3.3.3)
DE (Coproporphyrinogenase) (Coprogen oxidase).
GN HEMF OR SL1185
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,
RA Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
RA Yamada M., Yasuda M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -!- CATALYTIC ACTIVITY: Coproporphyrinogen-III + O(2) =
CC protoporphyrinogen-IX + 2 CO(2).
CC -!- COFACTOR: IRON (BY SIMILARITY).
CC -!- PATHWAY: Porphyrin biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE AEROBIC COPROPORPHYRINOGEN III OXIDASE
CC FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; D90901; BAA16863.1; -;
DR InterPro; IPR001260; Coprogen_oxidase.
DR Pfam; PF01218; Coprogen_oxidase; 1.
DR PRINTS; PR00073; COPROGNOXDASE.
DR PROSITE; PS01021; COPROGEN OXIDASE; 1.
KW Porphyrin biosynthesis; Oxidoreductase; Iron; Complete proteome.
SQ SEQUENCE 340 AA; 38937 MW; 9156238D89587FD CRC64;

Query Match 43.5%; Score 47; DB 1; Length 340;
Best Local Similarity 55.6%; Pred. No. 11;
Matches 10; Conservative 3; Mismatches 1; Indels 4; Gaps 1;

OY 2 LPPEMSGSLKRRPQEG 19
I I I I I I I I I I I I
Db 87 LPP-----SILKORPEAEG 100

RESULT 8

VP19_CNV
ID VP19_CNV STANDARD; PRT; 173 AA.

PI5184;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-AUG-1990 (Rel. 15, Last annotation update)
DE Core protein P19 (P20).
OS Cucurbit necrosis virus (CNV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tombusviridae;
OC Tombusvirus.
OX NCBI_TaxID=12143;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=89204896; PubMed=2705296;
RA Rochon D.M., Tremaine J.H.;
RT "Complete nucleotide sequence of the cucumber necrosis virus genome.";
RL Virology 169:251-259(1989).
CC -!- SIMILARITY: TO OTHER TOMBUSVIRUSES CORE PROTEIN P19.
CC -----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; M25270; AAA42906.1; -;
DR PIR; JA0133; NKVGCU.
DR InterPro; IPR004905; Tombus_P19.
DR Pfam; PF03220; Tombus_P19; 1.
KW Core protein.
SQ SEQUENCE 173 AA; 19808 MW; D05F870D4AA4C80F CRC64;
Query Match 42.6%; Score 46; DB 1; Length 173;
Best Local Similarity 53.3%; Pred. No. 7.2;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 3 PPEMSGSLKRRPQ 17
I I I I I I I I I I I I
Db 149 PCEVEGNVRRRPOG 163

RESULT 9

JAK3_HUMAN
ID JAK3_HUMAN STANDARD; PRT; 1124 AA.
AC P52333; Q13259; Q13260; Q13611;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase JAK3 (EC 2.7.1.112) (Janus kinase 3) (JAK-3)
GN JAK3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=94294384; PubMed=802790;
RA Kawamura M., McVicar D.W., Johnston J.A., Blake T.B., Chen Y.-Q.,
RA Lai B.K., Lloyd A.R., Kelvin D.J., Staples J.E., Ortaldo J.R.,
RA O'Shea J.J.;
RT "Molecular cloning of L-JAK, a Janus family protein-tyrosine kinase
RT expressed in natural killer cells and activated leukocytes.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:6374-6378(1994).
RN [2]
SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).
RX MEDLINE=96027605; PubMed=7559633;
RA Lai K.S., Jin Y., Graham D.K., Witthuhn B.A., Ihle J.N., Liu E.T.;
RT "A kinase-deficient splice variant of the human JAK3 is expressed in
RT hematopoietic and epithelial cancer cells.";
RL J. Biol. Chem. 270:25028-25036(1995).
RN [3]
SEQUENCE OF 36-191 FROM N.A.

MEDLINE=96278845; PubMed=8662778;
Verbsky J.W., Bach E.A., Fang Y.F., Yang L., Randolph D.A.,
Fields L.E.;
"Expression of Janus kinase 3 in human endothelial and other non-
lymphoid and non-myeloid cells.";
J. Biol. Chem. 271:13976-13980(1996).
[4]
VARIANT SCID CYS-100.
MEDLINE=95368142; PubMed=7659163;
Macchi P., Villa A., Gillani S., Sacco M.G., Frattini A., Porta F.,
Ugazio A.G., Johnston J.A., Candotti F., O'Shea J.J., Vezzoni P.,
Notarangelo L.D.;
"Mutations of Jak-3 gene in patients with autosomal severe combined
immune deficiency (SCID).";
Nature 377:65-68(1995).
[5]
VARIANTS SCID GLY-481; 586-LEU--MET-592 DEL AND ARG-759.
MEDLINE=98022793; PubMed=9354668;
Candotti F., Oakes S., Johnston J.A., Gillani S., Schumacher R.F.,
Mella P., Fiorini M., Ugazio A.G., Badolato R., Notarangelo L.D.,
Bozzi F., Macchi P., Strina D., Vezzoni P., Blaese R.M., O'Shea J.J.,
Villa A.;
"Structural and functional basis for JAK3-deficient severe combined
immunodeficiency.";
Blood 90:3996-4003(1997).
[6]
VARIANT SCID TRP-582.
MEDLINE=98423994; PubMed=9753072;
Bozzi F., Lefranc G., Villa A., Badolato R., Schumacher R.F.,
Khalil G., Loiselet J., Bresciani S., O'Shea J.J., Vezzoni P.,
Notarangelo L.D., Candotti F.;
"Molecular and biochemical characterization of JAK3 deficiency in a
patient with severe combined immunodeficiency over 20 years after
bone marrow transplantation: implications for treatment.";
Br. J. Haematol. 102:1363-1366(1998).
[7]
VARIANTS SCID ARG-151; ILE-722 AND SER-910.
MEDLINE=20435064; PubMed=10982185;
Schumacher R.F., Mella P., Badolato R., Fiorini M., Savoldi G.,
Gillani S., Villa A., Candotti F., Tampalini A., O'Shea J.J.,
Notarangelo L.D.;
"Complete genomic organization of the human JAK3 gene and mutation
analysis in severe combined immunodeficiency by single-strand
conformation polymorphism.";
Hum. Genet. 106:773-79(2000).
[8]
FUNCTION: TYROSINE KINASE OF THE NON-RECEPTOR TYPE, INVOLVED IN
THE INTERLEUKIN-2 AND INTERLEUKIN-4 SIGNALING PATHWAY.
PHOSPHORYLATES STAT6, IRS1, IRS2 AND PI3K.
[9]
CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
tyrosine phosphate.
[10]
SUBCELLULAR LOCATION: WHOLLY INTRACELLULAR, POSSIBLY MEMBRANE
ASSOCIATED (BY SIMILARITY).
[11]
ALTERNATIVE PRODUCTS: 3 isoforms; 1/JAK3B/breast-JAK3,
2/JAK3S/spleen-JAK3 (shown here) and 3/JAK3M/activated monocytes-
JAK3; are produced by alternative splicing. Isoform 1 may be
defective as it lack some part of the kinase domain.
[12]
TISSUE SPECIFICITY: IN NK CELLS AND AN NK-LIKE CELL LINE BUT NOT
IN RESTING T CELLS OR IN OTHER TISSUES. THE S-FORM IS MORE
COMMONLY SEEN IN HEMATOPOIETIC LINES, WHEREAS THE B- AND M-FORMS
ARE DETECTED IN CELLS BOTH OF HEMATOPOIETIC AND EPITHELIAL
ORIGINS.
[13]
DOMAIN: POSSESSES TWO PHOSPHOTRANSFERASE DOMAINS. THE SECOND ONE
PROBABLY CONTAINS THE CATALYTIC DOMAIN (BY SIMILARITY), WHILE THE
PRESENCE OF SLIGHT DIFFERENCES SUGGEST A DIFFERENT ROLE FOR
DOMAIN 1.
[14]
PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IL-2 AND IL-4.
[15]
DISEASE: DEFECTS IN JAK3 ARE A CAUSE OF RECESSIVE T-CELL
NEGATIVE/B-CELL POSITIVE SEVERE COMBINED IMMUNODEFICIENCY (T-B+
SCID). A CONDITION CHARACTERIZED BY THE ABSENCE OF CIRCULATING
MATURE T-LYMPHOCYTES AND NK CELLS, NORMAL TO ELEVATED NUMBERS OF
NONFUNCTIONAL B-LYMPHOCYTES, AND MARKED HYPOPLASIA OF LYMPHOID
TISSUES.

FT	SIGNAL	1	30	POTENTIAL.
FT	CHAIN	31	415	TUMOR NECROSIS FACTOR RECEPTOR
FT				SUPERFAMILY MEMBER 3.
FT				EXTRACELLULAR (POTENTIAL).
FT	DOMAIN	31	223	POTENTIAL.
FT	TRANSMEM	224	244	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	245	415	
FT	REPEAT	42	81	TNFR-CYS 1.
FT	REPEAT	82	124	TNFR-CYS 2.
FT	REPEAT	125	170	TNFR-CYS 3.
FT	REPEAT	171	213	TNFR-CYS 4.
FT	DISULFID	43	58	BY SIMILARITY.
FT	DISULFID	59	72	BY SIMILARITY.
FT	DISULFID	62	80	BY SIMILARITY.
FT	DISULFID	83	98	BY SIMILARITY.
FT	DISULFID	101	116	BY SIMILARITY.
FT	DISULFID	104	124	BY SIMILARITY.
FT	DISULFID	126	132	BY SIMILARITY.
FT	DISULFID	139	150	BY SIMILARITY.
FT	DISULFID	142	169	BY SIMILARITY.
FT	DISULFID	172	187	BY SIMILARITY.
FT	CARBOHYD	40	40	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	179	179	N-LINKED (GLCNAC. . .) (POTENTIAL).
SEQ	SEQUENCE	415 AA;	44956 MW;	29B326A56AEF661 CRC64;

Query Match 41.7%; Score 45; DB 1; Length 415;
Best Local Similarity 72.7%; Pred. No. 26;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps

QY	8	GSLKKRRPGE	18	
		:	:	
DB	255	GTLKKRHPGE	265	

RESULT 11
DXR-ARATH STANDARD; PRT; 477 AA.

ID	DXR-ARATH	STANDARD;	PRT;	477 AA.
AC	Q9XFS9; Q9M6U2;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	1-deoxy-D-xylulose 5-phosphate reductoisomerase, chloroplast precursor			
DE	(EC 1.1.1.267) (DXP reductoisomerase) (1-deoxyxylulose-5-phosphate			
DE	reductoisomerase)			
GN	DXR OR AT5G62790 OR MQB2.11.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosid 1; Brassicales; Brassicaceae; Arabidopsids.			
NCBI_TaxID	3702;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Columbia;			
RA	Campos N., Lols L.M., Cunillera N., Carretero L., Ahumada I.,			
RA	Hoefler J.-F., Pale-Grosdemange C., Rohmer M., Ferrer A., Boronot A.,			
RT	"Isolation and characterization of a cDNA from Arabidopsis thaliana			
RT	encoding 1-deoxy-D-xylulose 5-phosphate reductoisomerase, the first			
RT	committed enzyme of the non-mevalonate pathway for isoprenoid			
RT	biosynthesis.";			
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.			
RP	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Columbia;			
RA	MEDLINE=98290546; PubMed=9628582;			
RA	Sato S., Kaneko T., Kotani H., Nakamura Y., Asamizu E., Miyajima N.,			
RA	Tabata S.;			
RT	"Structural analysis of Arabidopsis thaliana chromosome 5. IV.			
RT	Sequence features of the regions of 1,456,315 bp covered by nineteen			
RT	physically assigned pl and TAC clones.";			
RL	DNA Res. 5:41-54(1998).			
RP	[3]			
RP	SEQUENCE OF 72-477 FROM N.A.			
RC	STRAIN=cv. Columbia;			
RA	Schweiner J., Mueller C., Zeidler J., Lichtenthaler H.K.;			

DR Pfam; PF02874; ATP-synt_ab_N; 1.
 DR TIGRfams; TIGR01043; ATP_syn_A_arch; 1.
 DR PROSITE; PS00152; ATPase.ALPHA.BETA; 1.
 KW Hydrolase; ATP synthesis; Hydrogen ion transport; ATP-binding;
 KW Complete proteome. 240 ATP (POTENTIAL).
 FT NP-BIND 233
 SQ SEQUENCE 584 AA; 64894 MW; 90F01BD38B3D6D78 CRC64;

Query Match 40.7%; Score 44; DB 1; Length 584;
 Best Local Similarity 47.1%; Pred. No. 52;
 Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 LPPMSGSLKRRPQGE 18
 : : : : :
 Db 151 IPPNVEGKLTWIA PQGE 167

RESULT 14
 BAT3_HUMAN STANDARD; PRT; 1132 AA.
 ID BAT3_HUMAN
 AC P46379;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Large proline-rich protein BAT3 (HLA-B-associated transcript 3).
 GN BAT3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=T-cell;
 RX MEDLINE=90192810; PubMed=2156268;
 RA Banerji J., Sands J., Strominger J.L., Spies T.;
 RT "A gene pair from the human major histocompatibility complex encodes
 RT large proline-rich proteins with multiple repeated motifs and a
 RT single ubiquitin-like domain."
 RL Proc. Natl. Acad. Sci. U.S.A. 87:2374-2378(1990).
 CC 1- FUNCTION: UNKNOWN
 CC 1- SIMILARITY: CONTAINS 1 UBIQUITIN-LIKE DOMAIN.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M33519; AAA35587.1; -
 CC DR EMBL; M33521; AAA35588.1; -
 CC DR EMBL; M33520; AAA35588.1; JOINED.
 CC DR HSP; P02248; IUBI.
 CC DR Genew; HGNC:13919; BAT3.
 CC DR MIM; 142590; -
 CC DR InterPro; IPR000626; Ubiquitin.
 CC DR Pfam; PF00240; ubiquitin; 1.
 CC DR SMART; SM00213; UBQ; 1.
 CC DR PROSITE; PS00299; UBIQUITIN_1; 1.
 CC DR PROSITE; PS50053; UBIQUITIN_2; 1.
 CC Repeat.
 CC KW DOMAIN 17 77 UBIQUITIN-LIKE.
 CC FT DOMAIN 202 207 POLY-PRO.
 CC FT DOMAIN 242 636 4 X 29 AA APPROXIMATE REPEATS.
 CC FT REPEAT 242 270 1.
 CC FT REPEAT 415 443 2.
 CC FT REPEAT 574 602 3.
 CC FT REPEAT 608 636 4.
 CC FT DOMAIN 657 670 POLY-PRO.
 CC SQ SEQUENCE 1132 AA; 119504 MW; E28CA8A78C38DD18 CRC64;

Query Match 40.7%; Score 44; DB 1; Length 1132;

Best Local Similarity 50.0%; Pred. No. 1e+02;
 Matches 13; Conservative 1; Mismatches 6; Indels 6; Gaps 2;
 QY 1 PLPP-----EMSGSLKRRP--QGE GP 20
 : : : : :
 Db 1034 PQPPLSDAYLSGMPAKRRKTKMQGEGP 1059

RESULT 15
 VNUA_PVKA STANDARD; PRT; 1733 AA.
 ID VNUA_PVKA
 AC P33485;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE Probable nuclear antigen.
 OS Pseudorabies virus (strain Kaplan) (PRV).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=33703;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91021039; PubMed=2171211;
 RA Vitek C., Kozmik Z., Paces V., Schirm S., Schwyzer M.;
 RT "Pseudorabies virus immediate-early gene overlaps with an oppositely
 RT oriented open reading frame: characterization of their promoter and
 RT enhancer regions".
 RL Virology 179:365-377(1990).
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M34651; AAA47471.1; -
 CC DR PIR; B45344; B45344.
 CC FT DOMAIN 112 117 POLY-THR.
 CC FT DOMAIN 179 1733 GLY-RICH.
 CC FT DOMAIN 192 196 POLY-SER.
 CC FT DOMAIN 271 298 POLY-PRO.
 CC FT DOMAIN 304 308 POLY-ARG.
 CC FT DOMAIN 883 889 POLY-GLY.
 CC FT DOMAIN 1398 1405 POLY-GLY.
 CC SQ SEQUENCE 1733 AA; 172166 MW; 0C8CD8BE475BB5E2 CRC64;

Query Match 40.3%; Score 43.5; DB 1; Length 1733;
 Best Local Similarity 50.0%; Pred. No. 1.9e+02;
 Matches 10; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

QY 1 PLPPMSGSLKRRPQGE GP 20
 : : : : :
 Db 294 PQPPAGGS-ARRRRGGGP 312

Search completed: April 8, 2003, 10:49:43
 Job time : 4.38425 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2003, 10:46:30 ; Search time 2.52983 Seconds
(without alignments) 760.006 Million cell updates/sec

Title: US-09-917-372-1_COPY_216_235
Perfect score: 108
Sequence: 1 PLPPMSGSLKRRPQGECP 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80	74.1	435	I54182	tumor necrosis fac
2	51	47.2	897	S02032	alpha-actinin 2, s
3	50	46.3	125	E72716	hypothetical prote
4	49	45.4	172	NKVGDD2	core protein p19 -
5	49	45.4	1460	1 EDBEIF	immediate-early pr
6	49	45.4	1870	2 S37671	MHC class III hist
7	49	45.4	1872	2 S36152	MHC class III hist
8	49	45.4	2142	2 B35098	MHC class III hist
9	47	43.5	147	2 T38396	small nuclear ribo
10	47	43.5	340	2 S74712	coproporphyrinogen
11	46.5	43.1	138	2 B84430	hypothetical prote
12	46.5	43.1	572	2 T45392	hypothetical prote
13	46	42.6	173	1 NKVGCU	core protein p20 -
14	46	42.6	471	2 A84741	probable myosins
15	46	42.6	501	2 S76563	hypothetical prote
16	46	42.6	528	2 T52092	DNA-binding protei
17	46	42.6	1124	2 A55747	L-JAK protein-tyro
18	45.5	42.1	363	2 A56940	integral membrane
19	45	41.7	132	2 AE2344	hypothetical prote
20	45	41.7	210	2 T36862	probable two-compo
21	45	41.7	583	2 E75529	probable peptide A
22	45	41.7	702	2 H90757	probable oxidoredu
23	45	41.7	702	2 F85621	probable oxidoredu
24	45	41.7	2938	2 T30249	cell proliferation
25	44.5	41.2	596	2 F84589	probable protein k
26	44.5	41.2	3198	2 A43426	collagen alpha 2 f
27	44	40.7	472	2 G71503	probable replicat
28	44	40.7	472	2 H81665	replicative DNA he
29	44	40.7	502	2 T03019	probable ribosomal

30	44	40.7	505	2 C98300	hypothetical prote
31	44	40.7	505	2 AD2983	hypothetical prote
32	44	40.7	584	2 G69227	ATP synthase, subu
33	44	40.7	803	2 F59433	RhGAP protein [lm
34	44	40.7	1132	2 A35098	MHC class III hist
35	43.5	40.3	1733	1 B45344	probable nuclear a
36	43.5	40.3	2327	2 T42630	aggreca - bovine
37	43	39.8	165	1 R3RT10	ribosomal protein
38	43	39.8	165	2 H59404	ribosomal protein
39	43	39.8	284	2 AD3275	beta-ureidopropion
40	43	39.8	497	2 B75218	glycerol-3-phospha
41	43	39.8	696	2 S74984	long-chain-fatty-a
42	43	39.8	781	2 A37956	sulfate permease I
43	43	39.8	894	1 FAHUA2	alpha-actinin 2 -
44	43	39.8	1213	2 A41724	limb deformity (ld
45	43	39.8	1531	2 T42218	slit-1 protein hom

ALIGNMENTS

RESULT 1

I54182
tumor necrosis factor receptor 2-related protein - human
C:Species: Homo sapiens (man)
C>Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 17-Mar-2000
C:Accession: I54182
R:Baens, M.; Chaffanet, M.; Cassiman, J.J.; Van den Berghe, H.; Marynen, P.
Genomics 16, 214-218, 1993
A:Title: Construction and evaluation of a hncDNA library of human 12p transcribed seq
A:Reference number: I54182; MUID:93252381; PMID:8486360
A:Accession: I54182
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-435 <RES>
A:Cross-references: GB:L04270; NID:G339761; PIDN:AAA36757.1; PID:G339762
C:Genetics:
A:Gene: GDB:LTBR
A:Cross-references: GDB:1230195; OMIM:600979
A:Map position: 12p13.3-12p13.1
C:Superfamily: tumor necrosis factor receptor type 1; NGF receptor repeat homology

Query Match 74.1%; Score 80; DB 2; Length 435;

Best Local Similarity 35.7%; Pred. No. 0.00037;

Matches 20; Conservative 0; Mismatches 0; Indels 36; Gaps 1;

QY 1 PLPPMS-----GSLKRRPQGECP 20

DB 216 PLPPMSGTMLVLPLAFFLLATVFCIWKSHPSLCRKLGLSLKRRPQGECP 271

RESULT 2

S02032
alpha-actinin 2, skeletal muscle splice form SK - chicken
C:Species: Gallus gallus (chicken)
C>Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 13-Aug-1999
C:Accession: S02032; I50604; S36393; S27391; S15481
R:Arimura, C.; Suzuki, T.; Yanagisawa, M.; Imamura, M.; Hamada, Y.; Masaki, T.
Eur. J. Biochem. 177, 649-655, 1988
A:Title: Primary structure of chicken skeletal muscle and fibroblast alpha-actinins d
A:Reference number: S02032; MUID:89064821; PMID:3197725
A:Accession: S02032
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-897 <ARI>
A:Cross-references: GB:X13874; NID:G63787; PIDN:CAA32078.1; PID:G63788
A>Note: part of this sequence was confirmed by protein sequencing
R:Tokue, Y.; Goto, S.; Imamura, M.; Obinata, T.; Masaki, T.; Endo, T.
Exp. Cell Res. 197, 158-167, 1991

A:Title: Transfection of chicken skeletal muscle alpha-actinin cDNA into nonmuscle an

A:Reference number: I50604; MUID:92070385; PMID:1720388

A:Accession: I50604

A>Status: preliminary; translated from GB/EMBL/DDBJ

```

RESULT 4
NKVGD2
core protein p19 - Cymbidium ringspot virus
C:Species: Cymbidium ringspot virus
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 21-Jul-2000
C:Accession: JS0271; S05459
R:Grieco, F.; Burgvan, J.; Russo, M.
J. Gen. Virol. 70, 2533-2538, 1989
A:Title: Nucleotide sequence of the 3'-terminal region of Cymbidium ringspot virus RN
A:Reference number: A92801; MUID:89381709; PMID:2778443
A:Accession: JS0271
A:Molecule type: genomic RNA
A:Residues: 1-172 <GR1>
R:Grieco, F.; Burgvan, J.; Russo, M.
Nucleic Acids Res. 17, 6383, 1989
A:Title: The nucleotide sequence of Cymbidium ringspot virus RNA.
A:Reference number: JS0268; MUID:89366663; PMID:2771646
A:Accession: JS0271
A:Status: preliminary; translation not shown
A:Molecule type: genomic RNA
A:Residues: 1-172 <GR2>
A:Cross-references: EMBL:X15511; NID:g59020; PIDN:CAA33535.1; PID:g4469164
C:Comment: The genome is a single-stranded, positive-sense RNA.
C:Superfamily: tombusvirus core protein p19
C:Keywords: core protein

Query Match 45.4%; Score 49; DB 1; Length 172;
Best Local Similarity 47.1%; Pred. No. 5.8;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 PLPPMSGSLKKRRPQG 17
DB 147 PTPSEREGNVSRRRPEG 163

RESULT 5
EMBEF
immediate-early protein IE180 - suid herpesvirus 1 (strain Indiana-Funkhauser)
C:Species: suid herpesvirus 1
C:Date: 30-Jun-1990 #sequence_revision 30-Jun-1990 #text_change 16-Feb-1997
C:Accession: S04713
R:Cheung, A.K.
Nucleic Acids Res. 17, 4637-4646, 1989
A:Title: DNA nucleotide sequence analysis of the immediate-early gene of pseudorabies
A:Reference number: S04713; MUID:89315207; PMID:2546124
A:Accession: S04713
A:Molecule type: DNA
A:Residues: 1-1460 <CHE>
C:Superfamily: herpesvirus immediate-early protein IE175
C:Keywords: DNA binding; early protein; transcription regulation

Query Match 45.4%; Score 49; DB 1; Length 1460;
Best Local Similarity 45.0%; Pred. No. 53;
Matches 9; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 1 PLPPMSGSLKKRRPQGE 20
DB 180 PAPPAPPAPFAPRPGDGP 199

RESULT 6
S37671
MHC class III histocompatibility antigen HLA-B-associated protein 2 [similarity] - h
C:Species: Homo sapiens (man)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 15-Sep-2000
C:Accession: S37671
R:Bougueleret, L.
submitted to the EMBL Data Library, August 1992
A:Reference number: S37671
A:Accession: S37671
A:Status: preliminary
A:Molecule type: DNA

```


RESULT 9

T38396
small nuclear ribonucleoprotein associated protein b - fission yeast (Schizosaccharomycetes)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T38396
R:McLean, J.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
Submitted to the EMBL Data Library, February 1996

A:Reference number: Z21791
A:Accession: T38396
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-147 <MCL>
A:Cross-references: EMBL:Z69240; PIDN:CAA93231.1; GSPDB:GN00066; SPDB:SPAC36A3.08
A:Experimental source: strain 972h-; cosmid 26A3
C:Genetics:
A:Gene: SPDB:SPAC26A3.08
A:Map position: 1
A:Introns: 1/3

Query Match 43.5%; Score 47; DB 2; Length 147;
Best Local Similarity 47.8%; Pred. No. 9.7;
Matches 11; Conservative 1; Mismatches 7; Indels 4; Gaps 1;

Oy 1 PLPPEMSGSLK----RRPQGG 19
| | | | |
Db 85 PMDPSMRGSLLSGPGVAPAGRG 107

RESULT 10

S74712
coproporphyrinogen oxidase (EC 1.3.3.3) III - Synechocystis sp. (strain PCC 6803)
N:Alternate names: coproporphyrinogenase; protein slIII85
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, K.; Kuroki, Y.
DNA Res. 3, 109-136, 1996
A>Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp.

S.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S74712
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-340 <KAN>
A:Cross-references: EMBL:D90901; GB:AB001339; NID:g1651897; PIDN:BAAI6863.1; PID:g1651897
A>Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Gene: hemF
C:Superfamily: coproporphyrinogen oxidase
C:Keywords: oxidoeductase

Query Match 43.5%; Score 47; DB 2; Length 340;
Best Local Similarity 55.6%; Pred. No. 23;
Matches 10; Conservative 3; Mismatches 1; Indels 4; Gaps 1;

Oy 2 LPPEMSGSLKKRPQGG 19
||| ||| ||| |||
Db 87 LPP-----SILKORPEAEG 100

RESULT 11

B84430
hypothetical protein At2g01870 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: B84430
R:Liu, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Miao, J.; Wu, Y.; Zhang, L.; Shao, J.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, D.; Koo, H.; Moffat, K.S.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Ventresca, D.; Nierman, W.C.; White, N. 1999
Nature 402, 761-768,

LENGTH: 461
TYPE: PRT
ORGANISM: Homo sapiens
US-09-894-924-17

Query Match 15.58; Score 349.5; DB 10; Length 461;
Best Local Similarity 29.78; Pred. No. 3.1e-13;
Matches 113; Conservative 47; Mismatches 143; Indels 77; Gaps 22;

QY	9	APGLAMGPIVIGLIGELGLAASQDAVP-----PYASE-NOTCRDQEKYEYEPQHRICCSR	61
DB	2	APVAVMALAVGL-ELMAAA--HALPAQVAFPTPAPEGSTCR--LREYDQTAQMCCSK	56
QY	62	CPPGTYVSACSRIRDIYCATCAENSYNHNHYLFIQOLCR---PCDPVMGLEIAPCTS	118
DB	57	CSPGQHAQVFCITKTSIDYCDSEEDSTYQLMNMWPECLSCGRCSQDV---ETQACTR	112
QY	119	KRKTCCRCQPGMFCAMALE-CTHCELLSDCPPG-----TEAEKDEVGKGNHCVPCKA	172
DB	113	BQNRICTRPGMYCALSKQEGCRLCAPLRKCRPGFVARPGTETSIVV-----CKPCAP	166
QY	173	GHFONTSSPSARCOPHTCENGIVLEAAGTAQSDTTC--KNPLEPLPEMSGSL-----	225
DB	167	GTFSTSTSDICRPHQICN---VVALPGNASRDVACTSTSPTRSMAP---GAVHLPQ	219
QY	226	LKRRPQEGEPNPVAGSWEPKRAH-----PYFPDLVQPLPLISGQVS--PVS--TGLPAAP	276
DB	220	VSTRSQHTQPTP-----EPSTAPSTSFLLPMGPS--PPAEGSTGDFALPVLIVGTALG	272
QY	277	VLEAGVPO-----QOSPDLITREPOLE--PGEOSVAHGTNGIHV-----TGSMT	320
DB	273	LLIGVNVCVIMTQYKKKPLCLQREAKVPHLPADKARCTGPEQOHLITAPSSSSSSLE	332
QY	321	ITGNIYIYNGPVLGGPPGPG	340
DB	333	SSASALDRRAPTRNOQAPG	352

Search completed: April 8, 2003, 11:03:50
Job time : 45.852 secs

```

; PRIOR FILING DATE: 1995-11-09
; PRIOR APPLICATION NUMBER: 08/468,453
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/038,765
; PRIOR FILING DATE: 1993-03-13
; PRIOR APPLICATION NUMBER: 07/523,635
; PRIOR FILING DATE: 1990-05-10
; PRIOR APPLICATION NUMBER: 07/421,417
; PRIOR FILING DATE: 1989-10-13
; PRIOR APPLICATION NUMBER: 07/405,370
; PRIOR FILING DATE: 1989-09-11
; PRIOR APPLICATION NUMBER: 07/403,241
; PRIOR FILING DATE: 1989-09-05
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-758-124-2

Query Match          15.5%; Score 349.5; DB 10; Length 461;
Best Local Similarity 29.7%; Pred. No. 3,1e-13;
Matches 113; Conservative 47; Mismatches 143; Indels 77; Gaps 22;

QY 9 APGLAWGPLVLGLFGLLAASQPAVP-----PYASE-NQTCRDOEKEYEPQHRICCSR 61
DB 2 APVAVMALAVGL-ELMAAA--HALPAQVAFPTPYAPEPGSTCR--LREYYDQTAQMCCK 56

QY 62 CPPTGYVSAKCSRIIDYVATCAENSYNEHWNITLITQICR---PCDPVMGLEELAPCTS 118
DB 57 CSPGQHAKEYCTKTSIDYVDCSDCESTYTQIMNWPBCLSCGSCSSDQY---ETQACTR 112

QY 119 KRKTQCRQCPGMFCAMALE-CTHCELLSDCPG-----TEAEIKDEVGKGNHCVPCA 172
DB 113 EQNRICTCRGMWYCALSKQEGCRLCAPLRKCRPGFVARPGTETSDVY-----CKPCAP 166

QY 173 GHQONTSSPSARCOPHTRCENOGVLEAPGTASDTTC--KNPLEPLPEMSSGL----- 225
DB 167 GFSTNTTSSDTCRPHQICN---VVAIPGNASMDAVCTSTSPTRSMAP---GAVILPDP 219

QY 226 LKRRPQEGNPNVAGSWPEPKAH-----PYFPDLVQPLPLISGDVS-PVS--TGLPAP 276
DB 220 VSTRQOHQPTP-----EPSTAPSTSLPLMGPS--PRAEGSTGDFALPGLVGTALG 272

QY 277 VLEAGVPO-----QOSPDLITREPQLE--PGEOSQVAHGTNGIHV-----TGSMT 320
DB 273 ILIIGVNCVIMTYQKKRPLCLQREAKVPHLPADKARGTGGPEQOHLITAPSSSSSIE 332

QY 321 ITGNITYINGPVLAGPPGPG 340
DB 333 SSASALDRRAPTRNQPQAPG 352

RESULT 14
US-09-896-096A-17
; Sequence 17, Application US/09896096A
; Patent No. US20020061559A1
; GENERAL INFORMATION:
; APPLICANT: ASHKENAZI, AVI J
; APPLICANT: BOTSTEIN, DAVID
; APPLICANT: DODGE, KELLY H.
; APPLICANT: GURNEY, AUSTIN L.
; APPLICANT: KIM, KYUNG JIN
; APPLICANT: LAWRENCE, DAVID A.
; APPLICANT: PITTI, ROBERT
; APPLICANT: ROY, MARGARET A
; APPLICANT: TUMAS, DANIEL B
; APPLICANT: WOOD, WILLIAM I.
; TITLE OF INVENTION: DCR3 Polypeptide, A TNFR Homolog
; FILE REFERENCE: P11342 REVISED
; CURRENT APPLICATION NUMBER: US/09/896,096A
; CURRENT FILING DATE: 2001-06-28
```

```

; PRIOR APPLICATION NUMBER: US 09/157,289
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: US 60/059,288
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: US 60/094,640
; PRIOR FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 18
; SEQ ID NO 17
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-896-096A-17

Query Match          15.5%; Score 349.5; DB 10; Length 461;
Best Local Similarity 29.7%; Pred. No. 3,1e-13;
Matches 113; Conservative 47; Mismatches 143; Indels 77; Gaps 22;

QY 9 APGLAWGPLVLGLFGLLAASQPAVP-----PYASE-NQTCRDOEKEYEPQHRICCSR 61
DB 2 APVAVMALAVGL-ELMAAA--HALPAQVAFPTPYAPEPGSTCR--LREYYDQTAQMCCK 56

QY 62 CPPTGYVSAKCSRIIDYVATCAENSYNEHWNITLITQICR---PCDPVMGLEELAPCTS 118
DB 57 CSPGQHAKEYCTKTSIDYVDCSDCESTYTQIMNWPBCLSCGSCSSDQY---ETQACTR 112

QY 119 KRKTQCRQCPGMFCAMALE-CTHCELLSDCPG-----TEAEIKDEVGKGNHCVPCA 172
DB 113 EQNRICTCRGMWYCALSKQEGCRLCAPLRKCRPGFVARPGTETSDVY-----CKPCAP 166

QY 173 GHQONTSSPSARCOPHTRCENOGVLEAPGTASDTTC--KNPLEPLPEMSSGL----- 225
DB 167 GFSTNTTSSDTCRPHQICN---VVAIPGNASMDAVCTSTSPTRSMAP---GAVILPDP 219

QY 226 LKRRPQEGNPNVAGSWPEPKAH-----PYFPDLVQPLPLISGDVS-PVS--TGLPAP 276
DB 220 VSTRQOHQPTP-----EPSTAPSTSLPLMGPS--PRAEGSTGDFALPGLVGTALG 272

QY 277 VLEAGVPO-----QOSPDLITREPQLE--PGEOSQVAHGTNGIHV-----TGSMT 320
DB 273 ILIIGVNCVIMTYQKKRPLCLQREAKVPHLPADKARGTGGPEQOHLITAPSSSSSIE 332

QY 321 ITGNITYINGPVLAGPPGPG 340
DB 333 SSASALDRRAPTRNQPQAPG 352

RESULT 15
US-09-894-924-17
; Sequence 17, Application US/09894924
; Patent No. US20020065210A1
; GENERAL INFORMATION:
; APPLICANT: ASHKENAZI, AVI J
; APPLICANT: BOTSTEIN, DAVID
; APPLICANT: DODGE, KELLY H.
; APPLICANT: GURNEY, AUSTIN L.
; APPLICANT: KIM, KYUNG JIN
; APPLICANT: LAWRENCE, DAVID A.
; APPLICANT: PITTI, ROBERT
; APPLICANT: ROY, MARGARET A
; APPLICANT: TUMAS, DANIEL B
; APPLICANT: WOOD, WILLIAM I.
; TITLE OF INVENTION: DCR3 Polypeptide, A TNFR Homolog
; FILE REFERENCE: P11342 REVISED
; CURRENT APPLICATION NUMBER: US/09/894,924
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 09/157,289
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: US 60/059,288
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: US 60/094,640
; PRIOR FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 18
; SEQ ID NO 17
```

```

: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: BROWDY AND NEIMARK
: STREET: 419 Seventh Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20004
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/800,909
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/476,862
: FILING DATE:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: IL 94039
: FILING DATE: 06-APR-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: IL 91229
: FILING DATE: 06-AUG-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: IL 90339
: FILING DATE: 18-MAY-1989
: ATTORNEY/AGENT INFORMATION:
: NAME: BROWDY, Roger L.
: REGISTRATION NUMBER: 25,618
: REFERENCE/DOCKET NUMBER: WALLACH-12A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-737-3528
: TELEFAX: 202-737-3528
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 461 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-800-909-2

Query Match          15.5%; Score 349.5; DB 10; Length 461;
Best Local Similarity 29.7%; Pred. No. 3.1e-13;
Matches 113; Conservative 47; Mismatches 143; Indels 77; Gaps 22;

QY 9 ABGLAMGPLVLGLFGLAASOPQAVP-----PYASE-NQTCRQDEKEYEPQHRICSR 61
DB 2 AVVAWMAALAVGL-ELMAAA--HALPAQVAFPPVAPPEPSTCR--LREYYDOTAQMCCK 56
QY 62 CPGGYVSAKCSRINDYVATCAENSYNHNWYLTICQLCR---PCDPYMGLEETAPCTS 118
DB 57 CSPGQHAKEYCTKTDYVCDSCEDSTYTQLMWVPECLSCGSRCSDDV---ETQACTR 112
QY 119 KRKTCCRCOPGMFCAMALE-CTHCELSDDCPG-----TEAELDEYKGNHCVPCA 172
DB 113 EGNRICTCRPGWYCALSKNOGCRCLCAPLKKCRPGGVAPRGVETSDV-----CKPCAP 166
QY 173 GHFQNTSSPBARCOPHTRCENGLVEAAGTASDPTTC--KNPLPLPEMGSST----- 225
DB 167 GFSTNTSTSDICRPHQICN---VVAIPGNASMDAVCTSTSPTRSMAP---GAVHLPPQ 219
QY 226 LKRRPQGEPPNVASWEPRKAH-----PYFPDLVQPLPLPSGVVS-PVS--TGLPAAP 276
DB 220 VSTRSQHTQPTP-----EPSTAPSTSFLLPMGPS--PPAEGSTGDFALFVGLIVGTALG 272
QY 277 VLEAGVPO-----QOSPDLTTRBPOLR--PGEOSOVAHGTNGIHW-----TGGSMT 320
DB 273 LLITGVNCVINTQYKKRKLCLQREAKVPHLPADAKRGTQGEQOHLITAPSSSSSILE 332
QY 321 ITGNITYINGPVLGGPPGPG 340

```

```

DB 333 SSASALDRRAPTRNQPOAPG 352

RESULT 12
US-09-826-212-4
: Sequence 4, Application US/09826212
: Patent No. US20010021516A1
: GENERAL INFORMATION:
: APPLICANT: Wei, Ying-Fei
: APPLICANT: Gentz, Reiner
: APPLICANT: Ruben, Steven
: APPLICANT: Ni, Jian
: TITLE OF INVENTION: Tumor Necrosis Factor Receptor 5
: FILE REFERENCE: 1488,1280006
: CURRENT APPLICATION NUMBER: US/09/826,212
: CURRENT FILING DATE: 2001-04-05
: NUMBER OF SEQ ID NOS: 26
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 4
: LENGTH: 461
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-826-212-4

Query Match          15.5%; Score 349.5; DB 10; Length 461;
Best Local Similarity 29.7%; Pred. No. 3.1e-13;
Matches 113; Conservative 47; Mismatches 143; Indels 77; Gaps 22;

QY 9 ABGLAMGPLVLGLFGLAASOPQAVP-----PYASE-NQTCRQDEKEYEPQHRICSR 61
DB 2 AVVAWMAALAVGL-ELMAAA--HALPAQVAFPPVAPPEPSTCR--LREYYDOTAQMCCK 56
QY 62 CPGGYVSAKCSRINDYVATCAENSYNHNWYLTICQLCR---PCDPYMGLEETAPCTS 118
DB 57 CSPGQHAKEYCTKTDYVCDSCEDSTYTQLMWVPECLSCGSRCSDDV---ETQACTR 112
QY 119 KRKTCCRCOPGMFCAMALE-CTHCELSDDCPG-----TEAELDEYKGNHCVPCA 172
DB 113 EGNRICTCRPGWYCALSKNOGCRCLCAPLKKCRPGGVAPRGVETSDV-----CKPCAP 166
QY 173 GHFQNTSSPBARCOPHTRCENGLVEAAGTASDPTTC--KNPLPLPEMGSST----- 225
DB 167 GFSTNTSTSDICRPHQICN---VVAIPGNASMDAVCTSTSPTRSMAP---GAVHLPPQ 219
QY 226 LKRRPQGEPPNVASWEPRKAH-----PYFPDLVQPLPLPSGVVS-PVS--TGLPAAP 276
DB 220 VSTRSQHTQPTP-----EPSTAPSTSFLLPMGPS--PPAEGSTGDFALFVGLIVGTALG 272
QY 277 VLEAGVPO-----QOSPDLTTRBPOLR--PGEOSOVAHGTNGIHW-----TGGSMT 320
DB 273 LLITGVNCVINTQYKKRKLCLQREAKVPHLPADAKRGTQGEQOHLITAPSSSSSILE 332
QY 321 ITGNITYINGPVLGGPPGPG 340
DB 333 SSASALDRRAPTRNQPOAPG 352

RESULT 13
US-09-758-124-2
: Sequence 2, Application US/09758124
: Patent No. US20020006391A1
: GENERAL INFORMATION:
: APPLICANT: SMITH, Craig A.
: APPLICANT: GOODWIN, Raymond G.
: TITLE OF INVENTION: TUMOR NECROSIS FACTOR-ALPHA AND -BETA RECEPTORS
: FILE REFERENCE: A7895
: CURRENT APPLICATION NUMBER: US/09/758,124
: CURRENT FILING DATE: 2001-01-12
: PRIOR APPLICATION NUMBER: 08/953,268
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 08/555,629

```

NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/112,793
FILING DATE: 28-Mar-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/828,683A
FILING DATE: 31-Mar-1997
APPLICATION NUMBER: 08/625328
FILING DATE: 1-Apr-1996
APPLICATION NUMBER: 08/710802
FILING DATE: 23-Sep-1996

ATTORNEY/AGENT INFORMATION:
NAME: Marschang, Diane L.
REGISTRATION NUMBER: 35,600
REFERENCE/DOCKET NUMBER: P1007P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-5416
TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 170 amino acids
TYPE: Amino Acid
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-10-112-793-14

Query Match 43.9%; Score 987; DB 9; Length 170;
Best Local Similarity 100.0%; Pred. No. 2e-50;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 TCRDQEKRYEPQHRICCSRCPPGYYSAKSRTIDYTCATCAENSYNEHNNYLTICQLC 101
DB 1 TCRDQEKRYEPQHRICCSRCPPGYYSAKSRTIDYTCATCAENSYNEHNNYLTICQLC 60

QY 102 RPPDVMGLELPIAFTSKRTQCRQCPMFCAMALECTHCELLSDCPGTEAELEKDEVG 161
DB 61 RPPDVMGLELPIAFTSKRTQCRQCPMFCAMALECTHCELLSDCPGTEAELEKDEVG 120

QY 162 KGNHCVCKAGHFQNTSSPSARCPHTRCENOGILEVAPGTAQSDTTC 211
DB 121 KGNHCVCKAGHFQNTSSPSARCPHTRCENOGILEVAPGTAQSDTTC 170

RESULT 10
US-09-768-779A-3
Sequence 3, Application US/09768779A
Patent No. US20020127637A1
GENERAL INFORMATION:
APPLICANT: NI, JIAN
TITLE OF INVENTION: HUMAN TUMOR NECROSIS FACTOR
RECEPTOR-LIKE PROTEIN 8
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: US
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/768,779A
FILING DATE: 25-Jan-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/086,582
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: KENLEY K. HOOVER
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PF368PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 450 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-768-779A-3

Query Match 15.5%; Score 349.5; DB 10; Length 450;
Best Local Similarity 29.7%; Pred. No. 3e-13;
Matches 113; Conservative 47; Mismatches 143; Indels 77; Gaps 22;

QY 9 APGLAWGPELVIGLFGLLASQPOAVP-----PYASR-NOTCRDQEKRYEPQHRICCSR 61
DB 2 APVAVMAALAVGL-ELWAAA--HALPAQVAFTPYAPPEPGSTCR--LREYYDOTAQMCCK 56

QY 62 CPGTYYSAKSRTIDYTCATCAENSYNEHNNYLTICQLC--PCDPVNGLEIACCTS 118
DB 57 CSPQDHAKVFCTKTSVDVCSCEJSTYTQLMNWPBELSCGSSRSQV-----ETQACTR 112

QY 119 KRKQCRQCPMFCAMALE-CTHCELLSDCPG-----TEAELEKDEVGKGNHCVCKA 172
DB 113 EQNNTICGRPGWCYALKSQECGRICAPLRKCRPGFGVAPGTERSDV-----CKPCAP 166

QY 173 GHFQNTSSPSARCPHTRCENOGILEVAPGTAQSDTTC--KNLEPLPPEMSSGL----- 225
DB 167 GFENNTSSPDICRPHQICN---VVALPGNSHDAVCTSTSPTRSMAP--GAVHLPP 219

QY 226 LKRPPGEGPMPVAGSMPEPKAH-----PYFPDLVQPLPIISGVS-PVS--TGLPAAP 276
DB 220 VSTRSQHTQPTP-----EPSTAPSTFLPPMGPS--PPABSGTDFAFLPVGLIYGVATG 272

QY 277 VLEAGVPQ-----QSPDLITREPQLE--PGQSOVAGHTNGIIV-----TGGSMT 320
DB 273 LLIGVNCVIMTQVKKKPLCLQREAKVPHLPADKARGTQGPQOHLITAPSSSSSSSLE 332

QY 321 ITGNIIYNGPVLAGPREG 340
DB 333 SSASALDRRAPTRNQPAAG 352

RESULT 11
US-09-800-909-2
Sequence 2, Application US/09800909
Patent No. US20010019833A1
GENERAL INFORMATION:
APPLICANT: WALLACH, David
APPLICANT: BIGDA, Jacek
APPLICANT: BELESKY, Igor
APPLICANT: METT, Igor
APPLICANT: ENGELMANN, Hartmut
TITLE OF INVENTION: TNF INHIBITORS

;; PRIOR APPLICATION NUMBER: 60/121,774
;; PRIOR FILING DATE: 1999-03-04
;; PRIOR APPLICATION NUMBER: 09/006,352
;; PRIOR FILING DATE: 1998-01-13
;; PRIOR APPLICATION NUMBER: 60/035,496
;; PRIOR FILING DATE: 1997-01-14
;; NUMBER OF SEQ ID NOS: 42
;; SOFTWARE: Patent In Ver. 2.1
;; SEQ ID NO: 8
;; LENGTH: 415
;; TYPE: PR
;; ORGANISM: Homo sapiens
US-09-935-727-8

Query Match 62.5%; Score 1404; DB 10; Length 415;
Best Local Similarity 63.9%; Pred. No. 4.3e-74;
Matches 265; Conservative 28; Mismatches 86; Indels 36; Gaps 4;

OY 1 MLPMATAPGLAMGLVGLGLGLAASQAPVPPASENOTCRDOEKEYEPOHRIICS 60
DB 1 MRLPRASSPCGLAMGLLGLGLGLVASQPOLVPPYRIENGTOMDQREYEPMDVCS 60
OY 61 RCPRETVYSAKSRIRDTVCATCAENSYNEHMYITICQLRCPDVPVGLIEIAPCTSKR 120
DB 61 RCPPEEFVAVCSRSQDVTCTCPHNSYNEHMYITICQLRCPDVIYGLFEVAPCTSDR 120
OY 121 KTCRCQGMFCAMALECTHC--ELLSDCPGTEAEIKDEYKGNHCVCKAGHPONT 178
DB 121 KAECCQGMSCYIYDNECVHCEERLVLCQPTAEVTDLMIDVAVCPCKPCHFORT 180
OY 179 SSPSARCCPHRCENOGVLEAAPTASQDTCKNLEP----- 216
DB 181 SSPRACCPHRCETCQGLVEAAPTASQDTCKNPEPAMLLALLSLVLELFTTVL 240
OY 217 -----LPEMS--GSLKRRPGEGRPVAGSWEPKAPHPDVLQPLDISGVSPV 268
DB 241 ACAMMRHPSLCKLSTLKRHEGESPPCPA---PRADPEPLAPLPMGDLSPS 296
OY 269 STGLPAFVLEAGVPOQSPDLITREPOLERPEQSOVAHGNGIHTVGGSMITITNITYI 328
DB 297 PAGPTAPSLLEVVLQOQSPVQARELEAPREHGVAGANGIHVTSVYVNTNITYI 356
OY 329 NGPVIGPPGPDLPATPEPPYPIPEEGDPGPGISTPHQEDGKAMHLETHCG 363
DB 357 NGPVIGTRGPDPPAPPEPPYPTPEEGAPGPESELSTYQEDGKAMHLETHCG 411

RESULT 7
US-09-948-018-19

;; Sequence 19, Application US/09948018
;; Patent No. US20020150977A1
;; GENERAL INFORMATION:
;; APPLICANT: Theell et al
;; TITLE OF INVENTION: TNF RECEPTOR-LIKE MOLECULES AND USES THEREOF
;; FILE REFERENCE: 0101737677
;; CURRENT APPLICATION NUMBER: US/09/948, 018
;; CURRENT FILING DATE: 2001-09-05
;; PRIOR APPLICATION NUMBER: US 60/230,191
;; PRIOR FILING DATE: 2000-09-05
;; NUMBER OF SEQ ID NOS: 45
;; SOFTWARE: Patent In version 3.1
;; SEQ ID NO: 19
;; LENGTH: 257
;; TYPE: PR
;; ORGANISM: Mus musculus
US-09-948-018-19

Query Match 53.9%; Score 1211; DB 10; Length 257;
Best Local Similarity 86.0%; Pred. No. 3.6e-63;
Matches 221; Conservative 0; Mismatches 0; Indels 36; Gaps 1;
OY 90 EHWNTLTICQLCRPDVPVGLIEIAPCTSKRKTQCRQGMFCAMALECTHCELLSDCP 149
|||||

DB 1 EHWNTLTICQLCRPDVPVGLIEIAPCTSKRKTQCRQGMFCAMALECTHCELLSDCP 60
OY 150 PGTAEELKDEYKGNHCVCKAGHPONTSSPSARCCPHRCENOGVLEAAPTASQDTT 209
DB 61 PGTAEELKDEYKGNHCVCKAGHPONTSSPSARCCPHRCENOGVLEAAPTASQDTT 120
OY 210 CKNPLEPLPEMS-----GSLKRRPGE 233
DB 121 CKNPLEPLPEMSGTMMLAVLLPLAFLLATVFCSCIMKSHPSLCKRLGSLKRRPGE 180
OY 234 GPNPVAGSWEPKAPHPYDVLQPLISGVSPVSTGLPAAPVLEAGVPOQSPDLITR 293
DB 181 GPNPVAGSWEPKAPHPYDVLQPLISGVSPVSTGLPAAPVLEAGVPOQSPDLITR 240
OY 294 EPOLERQSOVAHGNT 310
DB 241 EPOLERQSOVAHGNT 257

RESULT 8

US-10-003-211-1
;; Sequence 1, Application US/10003211
;; Publication No. US20020197254A1
;; GENERAL INFORMATION:
;; APPLICANT: Biogen, Inc.
;; TITLE OF INVENTION: Soluble Lymphotoxin Beta Receptor and
;; TITLE OF INVENTION: Anti-Lymphotoxin Receptor and Ligand Antibodies as
;; TITLE OF INVENTION: Therapeutic Agents for the treatment of Immunological
;; TITLE OF INVENTION: Diseases
;; FILE REFERENCE: A01305
;; CURRENT APPLICATION NUMBER: US/10/003, 211
;; CURRENT FILING DATE: 2001-10-31
;; PRIOR APPLICATION NUMBER: PCT/US97/19436
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/029, 060
;; PRIOR FILING DATE: 1996-10-25
;; NUMBER OF SEQ ID NOS: 1
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 1
;; LENGTH: 197
;; TYPE: PR
;; ORGANISM: Homo Sapien
US-10-003-211-1

Query Match 50.2%; Score 1129; DB 9; Length 197;
Best Local Similarity 99.5%; Pred. No. 1.5e-58;
Matches 196; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 28 SOPQAVPPYASENOTCRDOEKEYEPOHRIICSRCPGTYSYSAKSRIRDTVCATCAENS 87
DB 1 SOPQAVPPYASENOTCRDOEKEYEPOHRIICSRCPGTYSYSAKSRIRDTVCATCAENS 60
OY 88 YNEHMYITICQLCRPDVPVGLIEIAPCTSKRKTQCRQGMFCAMALECTHCELLSD 147
DB 61 YNEHMYITICQLCRPDVPVGLIEIAPCTSKRKTQCRQGMFCAMALECTHCELLSD 120
OY 148 CPGTEAEELKDEYKGNHCVCKAGHPONTSSPSARCCPHRCENOGVLEAAPTASQD 207
DB 121 CPGTEAEELKDEYKGNHCVCKAGHPONTSSPSARCCPHRCENOGVLEAAPTASQD 180
OY 208 TTCKNPLEPLPEMSGS 224
DB 181 TTCKNPLEPLPEMSGT 197

RESULT 9

US-10-112-793-14
;; Sequence 14, Application US/10112793
;; Publication No. US20020192729A1
;; GENERAL INFORMATION:
;; APPLICANT: Ashkenazi, Avi J.
;; TITLE OF INVENTION: Apo-2 LI AND Apo-3 POLYPEPTIDES

QY 325 IYINCPVLGGPPGDDLPATPEPPYPIPEEGDPGPGSLTPHOEDGKAMHLAETHECGA 384
 |||||
 Db 361 IYINCPVLGGPPGDDLPATPEPPYPIPEEGDPGPGSLTPHOEDGKAMHLAETHECGA 420
 |||||
 QY 385 TPNRGRPRNOFTIHD 399
 |||||
 Db 421 TPNRGRPRNOFTIHD 435

RESULT 4

US-09-826-212-6
 ; Sequence 6, Application us/09826212
 ; Patent No. US20010021516A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wei, Ying-Fei
 ; APPLICANT: Gentz, Reiner
 ; APPLICANT: Ruben, Steven
 ; APPLICANT: Ni, Jian
 ; TITLE OF INVENTION: Tumor Necrosis Factor Receptor 5
 ; FILE REFERENCE: 1488,1280006
 ; CURRENT APPLICATION NUMBER: US/09/826,212
 ; CURRENT FILING DATE: 2001-04-05
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 6
 ; LENGTH: 415
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-826-212-6

Query Match 62.5%; Score 1404; DB 10; Length 415;
 Best Local Similarity 63.9%; Pred. No. 4,3e-74;

Matches 265; Conservative 28; Mismatches 86; Indels 36; Gaps 4;

QY 1 MLIPWATSAPGLAMGPIVLGIFGLAASQPAVPYASENQTCRDOKEKEYEPOQRHCCS 60
 |||||
 Db 1 MLIPRASSPGGLAMGPIVLGIFGLAASQPAVPYRIENQTCWDQKEYEPEMHVCCS 60
 |||||
 QY 61 RCPGGTYVSAKSRIRDTVCATCAENSYNEHMYLTICQICRCPDPMGLIEIAPCTSKR 120
 |||||
 Db 61 RCPGGEVEFAVCSRSODTYCKTCPHNSYNEMHNLSTCQLCRPCDIYLGFEVEAPCTSDR 120
 |||||
 QY 121 KTCRCQPGMFCAMALECTHC--ELLSDCPGTEAELEKDEYKGNNGHCVCKAGHFONT 178
 |||||
 Db 121 KACRCQPGMSCVYLNECHCEERLYLQPGTEAEVDEIMDTVNCVCPKGFHONT 180
 |||||
 QY 179 SSPSARCQPHTRCENOGLEVAAPGTAOSDTTCNPLEP----- 216
 |||||
 Db 181 SSPSARCQPHTRCENOGLEVAAPGTAOSDTTCNPLEP----- 216
 |||||
 QY 217 -----LPEPMS-----GSLTKRRPQGECPNPVAGSMEPPKAHPYPPDIYOPLLPISGDVSPV 268
 |||||
 Db 241 ACAMMRHPSLCRKLGTILKRHPDEESPPCPA---PRADHPFDLAEPILPMSGDLSPS 296
 |||||
 QY 269 STGLPAAPVLEAGYVPOQSPDLITREPQLEPGESQVAHGNGIHYTGSMITGNIIYI 328
 |||||
 Db 297 PAGPPTASLEEVYVLOQSPVLQVARELEAEPGEGHGVANGIHVGGSVTVGNIIYI 356
 |||||
 QY 329 NGPVLGGPPGDDLPATPEPPYPIPEEGDPGPGSLTPHOEDGKAMHLAETHECG 383
 |||||
 Db 357 NGPVLGGPPGDDLPATPEPPYPIPEEGDPGPGSLTPHOEDGKAMHLAETHECG 411

RESULT 5

US-09-907-372-20
 ; Sequence 20, Application us/09907372
 ; Patent No. US20020068242A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lai, Preeti G.
 ; APPLICANT: Warren, Bridget A.
 ; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
 ; FILE REFERENCE: PC-0050 US
 ; CURRENT APPLICATION NUMBER: US/09/907,372

;; CURRENT FILING DATE: 2001-07-27
 ;; NUMBER OF SEQ ID NOS: 20
 ;; SOFTWARE: PERL Program
 ;; SEQ ID NO 20
 ;; LENGTH: 415
 ;; TYPE: PRT
 ;; ORGANISM: Mus musculus
 ;; FEATURE:
 ;; NAME/KEY: misc feature
 ;; OTHER INFORMATION: Incyte ID NO. US20020068242A1 9600223
 ; US-09-907-372-20

Query Match 62.5%; Score 1404; DB 10; Length 415;
 Best Local Similarity 63.9%; Pred. No. 4,3e-74;

Matches 265; Conservative 28; Mismatches 86; Indels 36; Gaps 4;

QY 1 MLIPWATSAPGLAMGPIVLGIFGLAASQPAVPYASENQTCRDOKEKEYEPOQRHCCS 60
 |||||
 Db 1 MLIPRASSPGGLAMGPIVLGIFGLAASQPAVPYRIENQTCWDQKEYEPEMHVCCS 60
 |||||
 QY 61 RCPGGTYVSAKSRIRDTVCATCAENSYNEHMYLTICQICRCPDPMGLIEIAPCTSKR 120
 |||||
 Db 61 RCPGGEVEFAVCSRSODTYCKTCPHNSYNEMHNLSTCQLCRPCDIYLGFEVEAPCTSDR 120
 |||||
 QY 121 KTCRCQPGMFCAMALECTHC--ELLSDCPGTEAELEKDEYKGNNGHCVCKAGHFONT 178
 |||||
 Db 121 KACRCQPGMSCVYLNECHCEERLYLQPGTEAEVDEIMDTVNCVCPKGFHONT 180
 |||||
 QY 179 SSPSARCQPHTRCENOGLEVAAPGTAOSDTTCNPLEP----- 216
 |||||
 Db 181 SSPSARCQPHTRCENOGLEVAAPGTAOSDTTCNPLEP----- 216
 |||||
 QY 217 -----LPEPMS-----GSLTKRRPQGECPNPVAGSMEPPKAHPYPPDIYOPLLPISGDVSPV 268
 |||||
 Db 241 ACAMMRHPSLCRKLGTILKRHPDEESPPCPA---PRADHPFDLAEPILPMSGDLSPS 296
 |||||
 QY 269 STGLPAAPVLEAGYVPOQSPDLITREPQLEPGESQVAHGNGIHYTGSMITGNIIYI 328
 |||||
 Db 297 PAGPPTASLEEVYVLOQSPVLQVARELEAEPGEGHGVANGIHVGGSVTVGNIIYI 356
 |||||
 QY 329 NGPVLGGPPGDDLPATPEPPYPIPEEGDPGPGSLTPHOEDGKAMHLAETHECG 383
 |||||
 Db 357 NGPVLGGPPGDDLPATPEPPYPIPEEGDPGPGSLTPHOEDGKAMHLAETHECG 411

RESULT 6

US-09-935-727-8
 ; Sequence 8, Application US/09935727
 ; Patent No. US20020150583A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Human Genome Sciences, Inc.
 ; TITLE OF INVENTION: Tumor Necrosis Factor Receptors 6 Alpha and 6 Beta
 ; FILE REFERENCE: PF454P2
 ; CURRENT APPLICATION NUMBER: US/09/935,727
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: 60/303,224
 ; PRIOR FILING DATE: 2001-07-06
 ; PRIOR APPLICATION NUMBER: 60/252,131
 ; PRIOR FILING DATE: 2000-11-21
 ; PRIOR APPLICATION NUMBER: 60/227,598
 ; PRIOR FILING DATE: 2000-08-25
 ; PRIOR APPLICATION NUMBER: 09/518,931
 ; PRIOR FILING DATE: 2000-03-03
 ; PRIOR APPLICATION NUMBER: 60/168,235
 ; PRIOR FILING DATE: 1999-12-01
 ; PRIOR APPLICATION NUMBER: 60/146,371
 ; PRIOR FILING DATE: 1999-08-02
 ; PRIOR APPLICATION NUMBER: 60/131,964
 ; PRIOR FILING DATE: 1999-04-30
 ; PRIOR APPLICATION NUMBER: 60/131,270
 ; PRIOR FILING DATE: 1999-04-27
 ; PRIOR APPLICATION NUMBER: 60/124,092
 ; PRIOR FILING DATE: 1999-03-12

[illegible]

```

RESULT 2
US-09-907-372-19
: Sequence 19, Application US/09907372
: Patent No. US20020068242A1
: GENERAL INFORMATION:
: APPLICANT: Lal, Preeti G.
: TITLE OF INVENTION: THE RECEPTOR 2 RELATED PROTEIN VARIANT
: FILE REFERENCE: PC-0050 US
: CURRENT APPLICATION NUMBER: US/09/907, 372
: CURRENT FILING DATE: 2001-07-27
: NUMBER OF SEQ ID NOS: 20
: SOFTWARE: PERL Program
: SEQ ID NO 19
: LENGTH: 435
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc.feature
: OTHER INFORMATION: Incyte ID No. US20020068242A1 g339762
US-09-907-372-19

```

Query Match	98.8%	Pred. 2230	DB 10	Length 435
Best Local Similarity	91.7%	Pred. No. 3.9e-121		
Matches 399	Conservative 0	Mismatches 0	Indels 36	Gaps 1

QY	1	MLLWMAISAPILANGPLVLGIFFGLILASQOAVPVPYASENOTCRDOKETEFYEPQHRCCS	60
Db	1	MLLPMWASAPILANGPLVLGIFFGLILASQOAVPVPYASENOTCRDOKETEFYEPQHRCCS	60
QY	61	RCPPGTIVYSAKCSIRDTVCATCAENSYNEHMNYLTICQLCRCPDPMGLLEETAPCTSKR	120
Db	61	RCPPGTIVYSAKCSIRDTVCATCAENSYNEHMNYLTICQLCRCPDPMGLLEETAPCTSKR	120
QY	121	KTCORCOPGMCAAMALECHTCELLSOPRGTAELEKDEYKGNHNCVPCAKGHQMTSS	180
Db	121	KTCORCOPGMCAAMALECHTCELLSOPRGTAELEKDEYKGNHNCVPCAKGHQMTSS	180
QY	181	PSARCOPTHRCENGGLVEAAPGTAAOSDTCKNPLLEAPRPMS	222
Db	181	PSARCOPTHRCENGGLVEAAGTAAOSDTCKNPLLEAPRPMS	240
QY	223	-----GSLIKRRPQGECPNPVAGSMEPPKAPHYRPDLVQPLPISGD	264
Db	241	ATVESCIIKSHPSLCKRLGISLILKRRPQGECPNPVAGSMEPPKAPHYRPDLVQPLPISGD	300
QY	265	VSPYSTGIPAPAVLEAGVPGQOQSPDLITRREQLERPGSQSOVAHGTNGIIHTGSGMTTGN	324
Db	301	VSPYSTGIPAPAVLEAGVPGQOQSPDLITRREQLERPGSQSOVAHGTNGIIHTGSGMTTGN	360
QY	325	IYIYNGPVLGSPGPGDLPATREPPYPIPEBGDGPBGGLSTPHQEDKAMHIAETEHGA	384
Db	361	IYIYNGPVLGSPGPGDLPATREPPYPIPEBGDGPBGGLSTPHQEDKAMHIAETEHGA	420
QY	385	TPSNRGPAPNOFITHD	399
Db	421	TPSNRGPAPNOFITHD	435

RESULT 3

US-09-768-779A-6
Sequence 6, Application US/09768779A
Patent No. US20020127637A1
GENERAL INFORMATION:
APPLICANT: NI, JIAN
MOORE, PAUL
TITLE OF INVENTION: HUMAN TUMOR NECROSIS FACTOR
RECEPTOR-LIKE PROTEIN 8
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/768, 779A
FILING DATE: 25-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/086,582
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: KENLEY K. HOOVER
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PF368BP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 435 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-768-779A-6

Query Match	98.8%	Score 2220	DB 10	Length 435		
Best Local Similarity	91.7%	Pred. No. 3.9e+121				
Matches 399	Conservative	0	Mismatches	0		
			Indels	36		
			Gaps	1		
Qy	1	MLLPATSA	PGLAMG	PVLG	LGFLGLAASQAPVAPYASENOTCRDOEKEYEPPQHRICCS	60
Db	1	MLLPATSA	PGLAMG	PVLG	LGFLGLAASQAPVAPYASENOTCRDOEKEYEPPQHRICCS	60
Qy	61	RCPPETIYSAKCS	RIRDPYCATCAENSTNEHNNYTLICOLCRPCDPVMGLIEIAC	TSK	120	
Db	61	RCPPETIYSAKCS	RIRDPYCATCAENSTNEHNNYTLICOLCRPCDPVMGLIEIAC	TSK	120	
Qy	121	KTCQCGQGMFC	AAALLECTHCELLSDCPGTEAEALKEVGGNNHCVCCKAGHQN	SS	180	
Db	121	KTCQCGQGMFC	AAALLECTHCELLSDCPGTEAEALKEVGGNNHCVCCKAGHQN	SS	180	
Qy	181	PSARQOPHTRC	ENQGLVEAPGTAOSDTCTCKNPLEPLRE	PMS	222	
Db	181	PSARQOPHTRC	ENQGLVEAPGTAOSDTCTCKNPLEPLRE	PMS	222	
Qy	223	-----GSLIKRRPQ	CGEGPNVAGSWEPKAAHYPFD	LVLPPLPISGD	264	
Db	241	ATVFSCIMKSHPS	LCRKLGSLIKRRPQCGEGPNVAGSWEPKAAHYPFD	LVLPPLPISGD	300	
Qy	265	VSPVSTGTPAA	AVLEAGVQOOSPLDLTRREPLEGEGSOVAHNGNGIHVYGGSWT	ITGN	324	
Db	301	VSPVSTGTPAA	AVLEAGVQOOSPLDLTRREPLEGEGSOVAHNGNGIHVYGGSWT	ITGN	360	

2

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2003, 10:51:46 ; Search time 42.852 Seconds
(without alignments)
569,244 Million cell updates/sec

Title: US-09-917-372-1

Perfect score: 2248
Sequence: 1 MLPWATSAPGLMGPVLV.....EHGATPSNRGPRNFITHD 399

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications, AA:*
1: /cgn2_6/prodata/1/pubpaa/us08_NEW_PUB pep:*
2: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB pep:*
3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB pep:*
4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB pep:*
5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB pep:*
6: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB pep:*
7: /cgn2_6/prodata/1/pubpaa/PCUS_PUBCOMB pep:*
8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB pep:*
9: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB pep:*
10: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB pep:*
11: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB pep:*
12: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB pep:*
13: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB pep:*
14: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2248	100.0	399	10	US-09-907-372-1
2	2220	99.8	435	10	US-09-907-372-19
3	2220	98.8	435	10	US-09-768-779A-6
4	1404	62.5	415	10	US-09-826-212-6
5	1404	62.5	415	10	US-09-907-372-20
6	1404	62.5	415	10	US-09-935-727-8
7	1211	53.9	257	10	US-09-948-018-19
8	1129	50.2	197	9	US-10-003-211-1
9	987	45.9	170	9	US-10-112-793-14
10	349.5	15.5	450	10	US-09-768-779A-3
11	349.5	15.5	461	10	US-09-800-909-2
12	349.5	15.5	461	10	US-09-826-212-4
13	349.5	15.5	461	10	US-09-758-124-2
14	349.5	15.5	461	10	US-09-896-096A-17
15	349.5	15.5	461	10	US-09-894-924-17
16	349.5	15.5	461	10	US-09-840-707A-17
17	349.5	15.5	461	10	US-09-800-908-3
18	349.5	15.5	461	10	US-09-935-727-6
19	349.5	15.5	461	12	US-10-164-592-3

20	327.5	14.6	474	10	US-09-758-124-4	Sequence 4, Appli
21	321.5	14.3	225	10	US-09-840-795-10	Sequence 10, Appl
22	319	14.2	300	9	US-10-129-709-3	Sequence 3, Appli
23	319	14.2	300	10	US-09-896-096A-1	Sequence 1, Appli
24	319	14.2	300	10	US-09-894-924-1	Sequence 2, Appli
25	319	14.2	300	10	US-09-935-727-2	Sequence 1, Appli
26	314.5	14.0	277	10	US-09-826-212-10	Sequence 10, Appl
27	314.5	14.0	277	10	US-09-768-779A-4	Sequence 4, Appli
28	314.5	14.0	277	10	US-09-935-727-12	Sequence 12, Appl
29	313.5	13.9	625	10	US-09-855-528-2	Sequence 2, Appli
30	312	13.9	625	9	US-09-877-650-15	Sequence 15, Appl
31	312	13.9	625	9	US-10-166-232A-2	Sequence 2, Appli
32	312	13.9	625	10	US-09-871-856-15	Sequence 15, Appl
33	311.5	13.9	277	10	US-09-839-339A-1	Sequence 1, Appli
34	310.5	13.8	299	10	US-09-877-156-17	Sequence 17, Appl
35	310	13.8	300	10	US-09-840-795-2	Sequence 2, Appli
36	309.5	13.8	235	9	US-10-243-230-2	Sequence 2, Appli
37	309.5	13.8	235	9	US-10-243-230-8	Sequence 8, Appli
38	309.5	13.8	235	9	US-09-882-735-16	Sequence 16, Appli
39	309.5	13.8	235	10	US-09-907-263-4	Sequence 4, Appli
40	308.5	13.7	235	9	US-10-243-230-4	Sequence 4, Appli
41	308.5	13.7	235	9	US-10-243-230-6	Sequence 6, Appli
42	306	13.6	227	10	US-09-840-795-9	Sequence 9, Appli
43	303.5	13.5	271	9	US-10-129-709-1	Sequence 1, Appli
44	303	13.5	184	9	US-09-852-455-8	Sequence 8, Appli
45	303	13.5	211	10	US-09-877-156-20	Sequence 20, Appli

ALIGNMENTS

RESULT 1
US-09-907-372-1
Sequence 1, Application US/09907372
Patent No. US20020068242A1
GENERAL INFORMATION:
APPLICANT: Lal, Preeti G.
TITLE OF INVENTION: TINE RECEPTOR 2 RELATED PROTEIN VARIANT
FILE REFERENCE: PC-0050 US
CURRENT FILING DATE: 2001-07-27
CURRENT APPLICATION NUMBER: US/09/907,372
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PERL Program
SEQ ID NO 1
LENGTH: 399
TYPE: PRT
ORGANISM: Homo sapiens
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No. US20020068242A1 7497867CD1
US-09-907-372-1

Query Match	100.0%	Score 2248	DB 10	Length 399
Best Local Similarity	100.0%	Pred. No. 8.8e-123		
Matches 399	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Db	1	MLPWATSAPGLMGPVLVIGFGLAASOPQAPVPAVSENOGCRDQKEYEYEQHRICCS	60	
Db	1	MLPWATSAPGLMGPVLVIGFGLAASOPQAPVPAVSENOGCRDQKEYEYEQHRICCS	60	
QY	61	RCPPGYVSANCSIRDTVCATCAENSYNEHWNLTTCQLCRPDDPVMGLEETAPCYSKR	120	
Db	61	RCPPGYVSANCSIRDTVCATCAENSYNEHWNLTTCQLCRPDDPVMGLEETAPCYSKR	120	
QY	121	KTCRCQCGMCAAMALECHTCHCLISDCCPPTTEALDEVGKGNHCVPCAGHFQNTSS	180	
Db	121	KTCRCQCGMCAAMALECHTCHCLISDCCPPTTEALDEVGKGNHCVPCAGHFQNTSS	180	
QY	181	PSARCQPTRENOGLVNAAPGTAQSDTTCNNPLEPPESSGSLKRRPQEGSPNVAG	240	
Db	181	PSARCQPTRENOGLVNAAPGTAQSDTTCNNPLEPPESSGSLKRRPQEGSPNVAG	240	

Matches 69; Conservative 22; Mismatches 83; Indels 60; Gaps 11;

OY 48 KEYTEPOHRIICSRCPQPTVYSACRSIRDTVCATCAENSYNEHMNYLTITQLCRPDDPV 107
| : | : | : | : | : | : | : | : | :
+-----+-----+

Dd 29 KHTVPSSGKH--CCRECOFGHGAVSRRCDIHRDILCHPCETGFGEVAANYNDT--CKQTQCNRH 86

OY 108 MGLEIAFCTSKRRTQQCGQGFCAMALECIECLSDCPPCTEALNDEVRKGNHC 167
- - - - - + + + + +
||| : ||| :

Dd 87 SSGSELKONCTPTODTVCCR-----PGTOR-PDSGYKLGVDC 123

OY 168 VPKAGHPQNFTSP-SARCOPIRCENOGLVEAPGAQSDTYCKMLPELPLEMSSSL 225
| | | | | : | : | : | : | : | : | : | : | :
+-----+-----+

Dd 124 VPCEPRGH---SGNNACKFPWTNCILTSKGOTHHPASDISIDANCED-----KSLLATL 173

OY 226 LKRKPQGGPNPVASWEPPRAHYEF-PDLVPLLPISGDVSYPSTGLPAAPVL 278

Db 174 I-----METOR--PTFRPTVO-----STWMPRTSELSPPPTL 205

RESULT 13

GQHDTI

tumor necrosis factor receptor 1 precursor [validated] - human

N:Alternate names: P55 tumor necrosis factor receptor; TNF receptor type 1

C:Contains: tumor necrosis factor alpha inhibitor; tumor necrosis factor binding protein

C:Species: Homo sapiens (man)

C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 08-Dec-2000

C:Accession: A38208, A34899, A34900, A36555, C36555; A38281; SI2057; JT0758; A60231; A38

R:Fuchs, P.; Strehl, S.; Dwozrak, M.; Hummler, A.; Ambros, P.F.

Genomics 13, 219-224, 1992

A>Title: Structure of the human TNF receptor 1 (p60) gene (TNRF1) and localization to ch

A:Reference number: A38208; MUTID:92250045; PMID:1315717

A:Accession: A38208

A:Molecule type: DNA

A:Residues: 1-455 <FUCC>

A:Cross-references: GB:M75864; GB:M75865; GB:M75866; NID:g339748; PID:NAA61201.1; PID:g

R:Lötscher, H.; Pan, Y.C.E.; Lahm, H.W.; Gentz, R.; Brochhaus, M.; Tabuchi, H.; Lesslau

Cell 61, 351-359, 1990

A>Title: Molecular cloning and expression of the human 55 kd tumor necrosis factor recep

A:Reference number: A34899; MUTID:90235284; PMID:2158862

A:Accession: A34899

A:Molecule type: mRNA

A:Residues: 1-455 <LOE>

A:Cross-references: GB:M58286; GB:M33480; NID:g339753; PID:NAAA36753.1; PID:g339754

A:Experimental source: Placenta

A>Note: Part of this sequence, including the amino end of the mature protein, confirmed

H:Schnall, T.J.; Lewis, M.; Köller, K.J.; Lee, A.; Rice, G.C.; Wong, G.H.W.; Galanaga, T

Cell 61, 361-370, 1990

A>Title: Molecular cloning and expression of a receptor for human tumor necrosis factor

A:Reference number: A34900; MUTID:90235285; PMID:2158863

A:Accession: A34900

A:Molecule type: mRNA

A:Residues: 1-455 <SCX>

A:Cross-references: GB:M3294; NID:g339744; PID:NAAA03210.1; PID:g339745

R:Hummel, A.; Maurer-Fogy, I.; Kroenke, M.; Scheutich, P.; Pfizenmaier, K.; Lantz, M.;

DNA Cell Biol. 9, 705-715, 1990

A>Title: Molecular cloning and expression of human and rat tumor necrosis factor recepto

A:Reference number: A36555; MUTID:91090841; PMID:1702293

A:Accession: A36555

A:Molecule type: mRNA

A:Residues: 1-455 <HNK>

A:Cross-references: GB:M63121; NID:g339755; PID:NAAA36754.1; PID:g339756

A:Accession: C36555

A:Molecule type: Protein

A:Residues: 30-38;41-53,'X','55'-79,'XX',82-94,'NK','XX',100-104;107-128;162-167,'X','169-262'

A>Note: The purified protein, called tumor necrosis factor binding protein, is a soluble

R:Gray, P.W.; Barrett, K.; Chanrity, D.; Turner, M.; Feldmann, M.

Proc. Natl. Acad. Sci. U.S.A. 87, 7380-7384, 1990

A>Title: Cloning of human tumor necrosis factor (TNF) receptor cDNA and expression of rec

A:Reference number: A38281; MUTID:91017509; PMID:2170974

A:Accession: A38281

A:Molecule type: mRNA

A:Residues: 1-455 <GRA>

A:Cross-references: GB:M37764

A:Note: the authors translated the codon TGG for residue 371 as Thr, AAG for residue Rbopnar, Y.; Kemper, O.; Brakebusch, C.; Engelmann, H.; Zwarg, R.; Aderka, D.; Holtm
 EMBO J. 9, 3269-3278, 1990
 A:Title: Soluble forms of tumor necrosis factor receptors (TNF-Rs). The cDNA for the
 le form of the receptor.
 A:Reference number: S12057; MUID:91006021; PMID:1698610
 A:Accession: S12057
 A:Molecule type: mRNA
 A:Residues: 1-455 <NO>
 A:Cross-references: EMBL:X55313; NID:937223; PIDD:CA939021.1; PID:937224
 A:Note: parts of soluble TNF binding protein 1, including its amino and carboxyl ends
 R:Kemper, O.; Wallach, D.
 Gene 134, 209-216, 1993
 A:Title: Cloning and partial characterization of the promoter for the human p55 tumor
 A:Reference number: J70758; MUID:94085779; PMID:8262379
 A:Accession: J70758
 A:Molecule type: DNA
 A:Residues: 1-13
 R:Seckinger, P.; Vey, E.; Turcatti, G.; Wingfield, P.; Dayer, J.M.
 Eur. J. Immunol. 20, 1167-1174, 1990
 A:Title: Tumor necrosis factor inhibitor: purification, NH-2-terminal amino acid seq
 A:Reference number: A60231; MUID:90292116; PMID:2113477
 A:Accession: A60231
 A:Molecule type: protein
 A:Residues: 41-43,'X',45-53,'X',55-57 <SRC>
 R:Galanaga, T.; Hwang, C.; Kohr, W.; Cappuccini, F.; Lucet III, J.A.; Jeffes, E.W.B.,
 Proc. Natl. Acad. Sci. U.S.A. 87, 8781-8784, 1990
 A:Title: Purification and characterization of an inhibitor (soluble tumor necrosis f
 clients.
 A:Reference number: A38258; MUID:91062364; PMID:2174164
 A:Accession: A38258
 A:Molecule type: protein
 A:Residues: 41-60 <GAT>
 A:Experimental source: cancer patient serum
 R:Rollson, I.; Lantz, M.; Nilsson, E.; Peetre, C.; Thyegell, H.; Grubb, A.; Adolf, G.
 Eur. J. Haematol. 42, 270-275, 1989
 A:Title: Isolation and characterization of a tumor necrosis factor binding protein f
 A:Reference number: A60594; MUID:89171156; PMID:2924890
 A:Accession: A60594
 A:Molecule type: protein
 A:Residues: 41-43,'X',45-53,'V',55-57,'XK',60 <OLS>
 A:Experimental source: renal failure patient urine
 R:Engelmann, H.; Novick, D.; Wallach, D.
 J. Biol. Chem. 265, 1531-1536, 1990
 A:Title: Two tumor necrosis factor-binding proteins purified from human urine. Evid
 A:Reference number: A35010; MUID:90110215; PMID:2153136
 A:Accession: A35010
 A:Molecule type: protein
 A:Residues: 41-45 <ENG>
 A:Experimental source: normal urine
 R:Kajihara, J.; Asada, A.; Kirihara, S.; Kato, K.
 Biosci. Biotechnol. Biochem. 58, 2266-2268, 1994
 A:Title: Amino acid sequence of natural tumor necrosis factor alpha inhibitor purifi
 A:Reference number: JC2404; MUID:95128033; PMID:7765720
 A:Accession: JC2404
 A:Molecule type: protein
 A:Residues: 41-53,'X',55-144,'X',146-150,'X',152-186,'X',188-201 <KAJ>
 A:Experimental source: urine
 C:Comment: This protein is one of two known receptors for both TNF-alpha (cachectin)
 C:Genetics:
 A:Gene: GDB:TNFR1
 A:Cross-references: GDB:125913; OMIM:191190
 A:Map position: 12p13.2-12p13.2
 A:Introns: 13/3: 65/1; 108/1: 158/1; 184/2: 209/1; 247/1: 256/3; 353/1
 C:Superfamily: tumor necrosis factor receptor type 1; NGF receptor repeat homology
 C:Keywords: duplication; glycoprotein; receptor; transmembrane protein
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-45/Prodot: tumor necrosis factor receptor 1 #status predicted <EXT>
 F:30-211/Domain: extracellular #status predicted <EXT>
 F:41-201/Prodot: TNF binding protein 1 (tumor necrosis factor alpha inhibitor) #sta
 F:44-82/Domain: NGF receptor repeat homology <NLI>
 F:84-126/Domain: NGF receptor repeat homology <NG3>
 F:127-167/Domain: NGF receptor repeat homology <NG2>

Db 86 SGESEKNCNPTEDTVQCR-----PSTQPR-QDSSKRLGVDC 122

QY 168 VPCKAGHPONTSSPASCQPTFRCENOGLYEAPGTAQSDTTCKNPLELPPEMGSLLK 227
 ||| : : : : : ||| : : : : :
 Db 123 VPCBPGRHSPGNOA--CKPWTNCTLSGKQIRHNASNSLDIVCD-----RSLLATIL- 173

QY 228 RRPQEGBNPVAWSPEPKAHPIYDPLIPISGVSFVSTGLPAAPVLEA 280
 ||| : : : : : ||| : : : : :
 Db 174 -----WETQRT-----TFRPTVPTSTVWPTSPQLSPRLPLVA 206

RESULT 11

CONST1

tumor necrosis factor receptor 1 precursor - mouse
 N:Alternate names: tumor necrosis factor receptor, 55K
 C:Species: Mus musculus (house mouse)
 C>Date: 30-Jun-1992 #sequence,revision 30-Jun-1992 #text,change 01-Dec-2000
 C:Accession: A38634; B40254; S16677; S19021; I54532; I57826
 R:Lewis, M.; Tartaglia, L.A.; Lee, A.; Bennett, G.L.; Rice, G.C.; Wong, G.H.W.; Chen, E.
 Proc. Natl. Acad. Sci. U.S.A. 88, 2830-2834, 1991

A:Title: Cloning and expression of cDNAs for two distinct murine tumor necrosis factor
 A:Reference number: A38634; MWID:91187885; PMID:1849278

A:Accession: A38634

A:Molecule type: mRNA

A:Residues: 1-454 <LEW>

A:Cross-references: GB:W60468; NID:q199825; PIDN:AAA39751.1; PID:q199826
 R:Goodwin, R.G.; Anderson, D.; Jerzy, R.; Davis, T.; Brannan, C.I.; Copeland, N.G.; Jenk
 Mol. Cell. Biol. 11, 3020-3026, 1991

A:Title: Molecular cloning and expression of the type 1 and type 2 murine receptors for
 A:Reference number: A40254; MWID:91246168; PMID:1645445

A:Accession: B40254

A:Molecule type: mRNA

A:Residues: 1-454 <GOZ>

A:Cross-references: GB:M60468; NID:q199825; PIDN:AAA39751.1; PID:q199826
 R:Barrett, K.; Taylor-Fishwick, D.A.; Cope, A.P.; Kissomergis, A.M.; Gray, P.W.; Feldma
 Eur. J. Immunol. 21, 1649-1656, 1991

A:Title: Cloning, expression and cross-linking analysis of the murine p55 tumor necrosis
 A:Reference number: S16677; MWID:91285014; PMID:1647956

A:Accession: S16677

A:Molecule type: mRNA

A:Residues: 1-454 <BAR>

A:Cross-references: EMBL:X59238; NID:g53578; PIDN:CAA1922.1; PID:g53579
 R:Rothe, J.G.; Brochhaus, M.; Gentz, R.; Lesslauer, W.

Immunogenetics 34, 338-340, 1991

A:Title: Molecular cloning and expression of the mouse Tnf receptor type b.
 A:Reference number: S19021; MWID:92039815; PMID:1657766

A:Accession: S19021

A:Molecule type: mRNA

A:Residues: 1-454 <ROT>

A:Cross-references: EMBL:X57796; NID:g54848; PIDN:CAA40936.1; PID:g54849
 R:Rebo, B.F.

Immunogenetics 39, 450-451, 1994

A:Title: Nucleotide sequence of the TNF type I receptor from a mouse endothelioma cell
 A:Reference number: I54532; MWID:94245292; PMID:8188324

A:Accession: I54532

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-454 <RES>

A:Cross-references: GB:L26349; NID:q430732; PIDN:AAA59361.1; PID:q430733
 R:Rothe, J.G.; Bluethmann, H.; Gentz, R.; Lesslauer, W.; Steinmetz, M.
 Mol. Immunol. 30, 165-176, 1993

A:Title: Genomic organization and promoter function of the murine tumor necrosis factor
 A:Reference number: I57826; MWID:93156721; PMID:8381516

A:Accession: I57826

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-393; 'G', 395-454 <RE2>

A:Cross-references: GB:M76656; NID:q202100; PIDN:AAA0465.1; PID:q202102
 C:Comment: This protein is one of two distantly related receptors for both TNF-alpha (ca
 C:Genetics: TNF-2

A:Gene: TNF-2

A:Introns: 13/3; 65/1; 108/1; 158/1; 184/2; 210/1; 248/1; 257/3; 353/1

C:Superfamily: tumor necrosis factor receptor type 1; NGF receptor repeat homology

C:Keywords: cytokine receptor; duplication; glycoprotein; receptor; transmembrane pr
 F:1-29/Domain: signal sequence #status predicted <SIG>
 F:30-454/Product: tumor necrosis factor receptor type 1 #status predicted <MAT>
 F:30-212/Domain: extracellular #status predicted <EXT>
 F:44-82/Domain: NGF receptor repeat homology <NG1>
 F:84-166/Domain: NGF receptor repeat homology <NG2>
 F:127-167/Domain: NGF receptor repeat homology <NG3>
 F:168-204/Domain: NGF receptor repeat homology <NG4>
 F:213-235/Domain: transmembrane #status predicted <MEM>
 F:236-454/Domain: intracellular #status predicted <INT>
 F:34,151,202/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.1%; Score 228; DB 1; Length 454;
 Best local similarity 25.4%; Pred. No. 3.7e-06;
 Matches 90; Conservative 40; Mismatches 149; Indels 76; Gaps 16;

QY 1 MLFWATSADELALGAPLVGLFGLAASQPAVP---PYASENQTCD---QEKEYEP 53
 ||| : : : : : ||| : : : : :
 Db 1 MGLP---TVPGIL---LSVLALLMGILHPSGVGLVPSLGDREK-RDSLCPOGKYVHSK 53

QY 54 QHRICSRCPGPGTVSAKC-SRIEDYCATCAENSVEHNNYLTICQLCRCPDVGLEE 112
 ||| : : : : : ||| : : : : :
 Db 54 NNSICTCKCHKGYLVSDCPSGCRDVCRCCEKGTFFAASQNYLQCLSCCKREMSQVE 113

QY 113 IAPCTSKRTQCRCPGPMCAAMALFECTHCEILSDCPGTEAELEKDEVGKGNHVCCKA 172
 ||| : : : : : ||| : : : : :
 Db 114 ISPCQAKDVCCCKENQF---QRLSTHITQCV-DCSPCNGVYTICKETQNTVCNCHA 170

QY 173 GHFQNTSS--PSARCQPHTRC-----ENQGLVEAP----- 201
 ||| : : : : : ||| : : : : :
 Db 171 GFFLRESECVPCSHCKKNECMKLCPLPPLANTNPODGTAVLLPLVILGLCLLSFIF 230

QY 202 -----GTQSDPTCKNPLELPPEMGSLLKRRPQ-----GEGRPVAGSWPEPK 246
 ||| : : : : : ||| : : : : :
 Db 231 ISLMCRTPRRRPEVYSITCDPV-PVKEEKAGKPLPAPSPAPSPISGFPPTLGFSTPGF 289

QY 247 AHFYFDLPVLPPLPISGD-----VSPYSTGLP---AAPVLEAGVPOQSPDLTR 293
 ||| : : : : : ||| : : : : :
 Db 290 SSPVSS---TPISPIEFGSPNMHMPVSEVVPVPGGAPPLLYESLCSVPATTSYQK 341

RESULT 12

148700

gene ox40 protein - mouse
 N:Alternate names: Ox40 antigen
 C:Species: Mus musculus (house mouse)
 C>Date: 02-Jul-1996 #sequence,revision 02-Jul-1996 #text,change 11-Jan-2000
 C:Accession: I48700; I48334; S34377
 R:Calderhead, D.M.; Buhlmann, J.E.; van den Eertwegh, A.J.; Claassen, E.; Noelle, R.J
 J. Immunol. 151, 5261-5271, 1993

A:Title: Cloning of mouse Ox40: a T cell activation marker that may mediate T-B cell
 A:Reference number: I48700; MWID:94044750; PMID:8228223

A:Accession: I48700

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-272 <RES>

A:Cross-references: EMBL:Z21674; NID:g312827; PIDN:CAA79772.1; PID:g312828
 R:Blirkeland, M.L.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.; Barclay, A.N.
 Eur. J. Immunol. 25, 926-930, 1995

A:Title: Gene structure and chromosomal localization of the mouse homologue of rat Ox
 A:Reference number: I48334; MWID:95255413; PMID:7737295

A:Accession: I48334

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-14; 'G', 16-272 <RE2>

A:Cross-references: EMBL:X85214; NID:g732818; PIDN:CAA59476.1; PID:g732819
 C:Genetics: Ox40

A:Gene: Ox40

A:Introns: 45/1; 86/1; 122/1; 144/2; 210/1; 250/1

C:Superfamily: CD27 antigen; NGF receptor repeat homology

Query Match 10.1%; Score 226; DB 2; Length 272;
 Best local similarity 29.5%; Pred. No. 2.9e-06;

QY 209 TCKNLEPLPMPSSGLKRRPOGEPNVPAGSWPPKAPHYEPDLVQPL----- 258
 DB 166 VCA-PESPILSAIPRTLYVSQPEPTRSQPL--DQEPGSGQT--PSILSLSGSTPIITEST 240
 QY 239 ----LPISGDVSPVSTGLPAAPVLEAG--VPQOSPLDLTREPOLE--PBGOSQVANGT 309
 DB 241 KGISLIPGLIYGVSTGLMGLVNCFILVQRKKKPSCLQGRDAVPHVPDEKSDAVGL 300
 QY 310 NGIHW 314
 DB 301 EQOHL 305

RESULT 6

death receptor-6 - chicken
 C:Species: Gallus gallus (chicken)
 C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
 C:Accession: J07705
 R:Bridgeham, J.T.; Bohe, J.; Goetz, F.W.; Johnson, A.L.
 Biochem. Biophys. Res. Commun. 284, 1109-1115, 2001
 A>Title: Conservation of death receptor-6 in avian and piscine vertebrates.
 A:Reference number: J07705; MUID:21308433; PMID:11414636
 A:Accession: J07705
 A:Molecule type: mRNA
 A:Residues: 1-651 <BRI>
 A:Comment: This receptor, a member of the tumor necrosis factor receptor family, belongs to the TNF receptor superfamily, and is involved in cell death and/or survival signaling cascade.
 C:Genetics:
 A:Gene: dr-6
 C:Keywords: ovary
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:52-196/Domain: extracellular cysteine-rich, ligand-binding #status predicted <ECL>
 F:332-335/Domain: transmembrane #status predicted <TM>
 F:410-475/Domain: death domain #status predicted <DED>
 F:551-651/Region: conserved cytoplasmic #status predicted

Query Match 11.9%; Score 267; DB 2; Length 651;
 Best Local Similarity 25.1%; Pred. No. 3.1e-08;

Matches 107; Conservative 46; Mismatches 150; Indels 124; Gaps 25;

QY 18 VLGLFGLLAASQPOAVPPVASENOTCRDQKEYE---POHRCSCRPCTGYSAKCS 73
 DB 9 VLPLVLFGLTADAQ--PKLTSEQNANVSLPAGKYIHDRATNGELICDKCPAGTYVSKHT 66
 QY 74 RIRDIYCATCAENSYNEHNNYLTICQLCR-PCDPVAGLEELIAPCTSKRTQCRQPGMF- 131
 DB 67 KSTLRCSGPCDGTFTKHNGIERCHPCRPCE--LPMTEKTHCTALMDRECTCLSGTPO 124
 QY 132 ----CAAMALECTHCELLSDCP-----PGTEAEIKDEVGKGNMNCVCKAGHQNNTSS 180
 DB 125 INDTCVPYIV-----CPVGCVCYRKKGTEYE-----DVACPKLKGTSDDVPS 166
 QY 181 PSARCQPHTRCENOGLEVAAPGTAOSDTYCKNPLEPLPMPSSGLKRRPOGEPNVPAG 240
 DB 167 SWMKCTYIDCGKKNVYVKKPGTKESDNCXSPAS-LP--NTSLSSAQAQDGETYEA- 221
 QY 241 SNEPKAPHYFP-DLVQPLPLISGDVSP-VSTGLPAAPVLEAGVPOQSPDLDTREPOLE 298
 DB 222 ----PPTA--YLPKGLNSSFVDFSSSPAPRVNSG-----TAEPTVD 256
 QY 299 PEGOSVANGTNGIHTGSMITGNI-----YIYNGPVLYGPPGP-----GDLPAIPE 347
 DB 257 YNDTS--ANGTVG--APGSLSSAGTAGAGSYRHKHTSQAMGKOPAOEMAGKESKIPY 311
 QY 348 PPYPIPEEDPDPGLSTPHODGKAMHLAETEH-----CGATPSNR 389
 DB 312 RP---PRR---GPPNV---HQ-----HPDINEHLPMLVLELLLVLYIVYVCSYAKSSR 356
 QY 390 ---GPR-392
 DB 390 ---GPR-392

DB 357 TLKKGPR 363

RESULT 7

A6476
 B cell-associated surface molecule CD40, long splice form - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 17-Nov-2000
 C:Accession: A46476; A46515
 R:Torres, R.M.; Clark, E.A.
 J. Immunol. 148, 620-626, 1992
 A>Title: Differential increase of an alternatively polyadenylated mRNA species of murine CD40.
 A:Reference number: A46476; MUID:92105763; PMID:1370315
 A:Accession: A46476
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-305 <TOR>
 A:Cross-references: GB:M83312; NID:g1553058
 A>Note: sequence extracted from NCBI backbone (NCBIN:75206, NCBI:75207)
 A>Note: this translation is not annotated in Genbank entry M83312, release 113.0
 R:Grimaldi, J.C.; Torres, R.; Kozak, C.A.; Chang, R.; Clark, E.A.; Howard, M.; Cockley
 J. Immunol. 149, 3921-3926, 1992
 A>Title: Genomic structure and chromosomal mapping of the murine CD40 gene.
 A:Reference number: A46515; MUID:93094586; PMID:1281194
 A:Accession: A46515
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: nucleic acid
 A:Residues: 1-287, 'lv' <GRI>
 A:Cross-references: GB:M83312; NID:g1553058; PIDN:AAB08705.1; PID:g1553059; GB:M94126
 A:Experimental source: BALB/c, liver
 A>Note: sequence extracted from NCBI backbone (NCBI:120357)
 C:Comment: For an alternative splice form, see PIR:A46515.
 C:Comment: For an alternative splice form, see PIR:A46476.
 C:Superfamily: CD27 antigen; NGF receptor repeat homology
 C:Keywords: alternative splicing; transmembrane protein
 F:105-144/Domain: NGF receptor repeat homology <NGF>

Query Match 11.6%; Score 260; DB 2; Length 305;
 Best Local Similarity 26.1%; Pred. No. 3.7e-08;

Matches 90; Conservative 38; Mismatches 137; Indels 80; Gaps 15;

QY 8 SAPGLA--WGPLVLGLFGLLAASQPOAVPPVASENOTCRDQKEYEPOHRCSCRP 65
 DB 3 SLPRICALWGCGL-----LTAVHLGCYV-----TCSQ--KOYLDHGO--CCDLCPG 44
 QY 66 TVYSAKCSIRDTYCATCAENSYNEHNNYLTICQLCRPCDPVAGLEELIAPCTSKRTQCR 125
 DB 45 SRTLSCTALEKTYQCHPCDSGEFSQAWNREIRCHQHRCEPNOGLRVKKEGTAESDYCT 104
 QY 126 COPMECAAMALECTHCELLSDCPG-----TEAEIKDEVGKGNMNCVCKAGHQNNTSS 180
 DB 105 CKEGQHCT--SKDCEACAGHTPCIFGFGVMEATETDTIV-----CHCPVGPFSNQS 156
 QY 181 PSARCQPHTRCENOGLEVAAPGTAOSDTTC--KNPLEPL--PMPSSGLK----- 227
 DB 157 LPEKCYPMTCSDKNLEVLQKTSOTNYICGLSKRMALLVLPVWGLITITFGFELYIK 216
 QY 228 ---RRQSGRPPVAGSWPPKAPHYFPDL-----VQPLPLISGDVSPVSTGL 272
 DB 217 KYVKKRPKDNEMLPAPARODPQEMEDYRGHNTAARVQETLHSCQPTQDEGKESRISV-- 274
 QY 273 PAAPVLEAGVPOQSPDL--LTFREPOLEPEGOSQVANGTNGIHW 315
 DB 275 -----QEROVTDISALRPGLNPG-----TAFGDDGLIT 304

RESULT 8

J04302
 tumor necrosis factor receptor p55 precursor - pig
 C:Species: Sus scrofa domestica (domestic pig)
 C>Date: 29-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 23-Jul-1999
 C:Accession: J04302; PC4093
 R:Suter, B.; Pauli, U.

A:Cross-references:	GDH:215268;	OMIM:109535
A:Map position:	20q12-20q13.2	
C:Superfamily:	CD27 antigen, NGE receptor repeat homology	
C:Keywords:	B-cell; glycoprotein; phosphoprotein; surface antigen; transmembrane protein	
F:1-30/Domain:	signal sequence #status predicted <SIG>	
F:21-277/Product:	B-cell activation protein CD40 #status experimental <MAT>	
F:21-193/Domain:	extracellular #status predicted <EXT>	
F:194-215/Domain:	transmembrane #status predicted <TM>	
F:216-277/Domain:	intracellular #status predicted <CYT>	
F:153,180/Binding site:	carbohydrate (Asn) (covalent) #status predicted	
Query Match	14.0%;	Score 314.5; DB:2;
Best Local Similarity	29.8%;	Pred. No. 2,6e-11;
Matches	92;	Conservative 38; Mismatches 120; Indels 59; Gaps 14;
QY	16	PVLGLFG-LLAASQPAVPPYASENOTCRDQEKYEYRPHRICSRCPETIYSAKCSR 74
DB	5	PLQCVLWGGCLLVAVHPE---PPTA-----CR-EKYLINSSQ---CCSLQPGQKIVSDTE 53

```

QY 75 IEDYICATCAENSYNEHHNNYLITQOLCRPDPMGLIEFIACSTKRRKQOCGOGMCAA 134
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 54 FTECEILPCGESEFLDTWNRRETHCHQHRYCDPNGLRVQOKGTSTDTITCEBGMHTS 113
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 135 MALECTHCELLSDCPPTGEAEELKDEVKGNHNCVCKAGHFOHTSSPSARCOPHTCENQ 194
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 114 EA--CESCVLHSCSPGGGVK-QIATGVSDTICECCPVGFTSNVSSANEKHPMTSCGTK 170
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 195 GLVEAPGTAOSDTTCKNPLEBLP-----PEMGSGL-----LKRPPGEGCNP 237
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 171 DLVQOAGNKTDDVVC-GPDRLRALVYPIIFGILFLVLVFEIKKVARKPNNKAPHP 229
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 238 VAGSMPEPPRAHVEFDLYQPLPIISGVDSPVSTGLPAPVLEAGVPOQOSPLDLTRPOL 297
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 230 ---KQEPQELIN--FPD-----DLPGSNTAPVQETLHGCPQYIQ----- 263
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 298 EPEQOSOVA 306
      | : : : : : :
Db 264 EDGKESRIS 272

```

A.Title: Allelic variation of the type 2 tumor necrosis factor receptor gene.
A.Reference number: I48854; MUID:95178848; PMID:7873884
A.Accession: I48854
A.Status: preliminary; translated from GR/EMBL/DDBJ
A.Molecule type: mRNA
A.Residues: 1-459 <RES>
A.Cross-references: EMBL.X76401; NTD:g433830; PIDN:CA553981.1; PTD:g433831
C.Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology F.I51-188/Domain: NGF receptor repeat homology <NGF>

Query Match 13.9%; Score 312; DB 2; Length 459;
Best Local Similarity 29.2%; Pred. No. 6e-11;
Matches 89; Conservative 41; Mismatches 135; Indels 40; Gaps 13;

DQ 35 PYASE-NOTCRDQEKEYEPORHICCSRCPPGTIVSACSRIDRTVCATCAENSNEYHN 93
 || : |::| : : |::|| : : |::|| : : |::|| : : |::|| : : |::||
DB 16 PKPREGVYCOQSQ-EYDRKQMOCCAPCPGQYVKHCNKNTSDTVCADCEASMYQTWVN 74

DQ 94 YLTTCQLCH---PCDPVMGLEIAPCTSKRKTCKRCOPGMCA--AMALECTHCELSDC 148
 || : |::| : : |::|| : : |::|| : : |::|| : : |::|| : : |::||
DB 75 QERTCLSCSSCSSTDOY----ETRACTKOQNNVACAGCAGRYCALKTHSGSRCOCRIKSLKC 130

DQ 149 PGTEAEIKDEGVKGNNHCVPCKASHFPONTSSPSARCOPHTRCGENGLVEAFPGAGSDT 208
 || : |::| : : |::|| : : |::|| : : |::|| : : |::|| : : |::||
DB 131 GGCF-GVASRAPNCNVLCACAPCTFSPTDSITVCSPHRICS---TLAIIPGNASTDA 185

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2003, 10:46:30 ; Search time 50.4702 Seconds

(without alignments)
760.006 Million cell updates/sec

Title: US-09-917-372-1

Perfect score: 2248

Sequence: 1 MLPMWATSAPGLAMGPLVLG.....EHCGATPNSNKGPNQFTTHD 399

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 segs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR-73:*
1: pirl:*
2: pirl2:*
3: pirl3:*
4: pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2220	98.8	435	2	154182	tumor necrosis fac
2	349.5	15.5	461	1	A35356	tumor necrosis fac
3	327.5	14.6	474	2	B38634	tumor necrosis fac
4	314.5	14.0	277	2	A60771	B-cell activation
5	312	13.9	459	2	I48854	gene murine tumour
6	267	11.9	651	2	JC7705	death receptor-6-
7	260	11.6	305	2	A46476	B cell-associated
8	244.5	10.9	461	2	JC4302	tumor necrosis fac
9	239	10.6	595	2	A42086	CD30 antigen precu
10	232	10.3	271	2	S12783	OX40 antigen precu
11	228	10.1	454	1	GOMST1	tumor necrosis fac
12	226	10.1	272	2	I48700	gene ox40 protein
13	222.5	9.9	455	1	GQHOT1	tumor necrosis fac
14	210	9.3	348	2	T28623	hypothetical prote
15	210	9.3	349	2	D36858	gene G4R protein -
16	209.5	9.3	277	2	I37552	OX40 homolog - hum
17	207	9.2	349	2	D72175	G2R protein - vari
18	204	9.1	461	1	GQRTT1	tumor necrosis fac
19	203.5	9.1	255	2	I38426	lymphocyte activat
20	196.5	8.7	425	1	A26431	nerve growth facto
21	191.5	8.5	493	2	JC5486	membrane glycoprot
22	189	8.4	325	2	B43692	T2 protein - rabbi
23	187.5	8.3	314	2	I37383	FAS soluble protei
24	185.5	8.3	256	2	B32393	T-cell antigen 4-1
25	183.5	8.2	326	1	GQVZML	T2 protein - myxom
26	183.5	8.2	335	2	A40036	apoptosis-mediatin
27	183.5	8.2	416	1	JN0006	nerve growth facto
28	175.5	7.8	427	1	GQHUN	nerve growth facto
29	162.5	7.2	327	2	A46484	apoptosis-mediatin

30	161.5	7.2	290	2	T24590	hypothetical prote
31	153	6.8	260	1	A46517	CD27 antigen precu
32	150	6.7	290	2	T24586	hypothetical prote
33	148.5	6.6	3635	2	T10053	laminin alpha 5 ch
34	148	6.6	303	2	T28999	hypothetical prote
35	146.5	6.5	1603	2	S23810	collagen alpha 1(X
36	145.5	6.5	324	2	JC2395	Fas antigen precu
37	145	6.5	1870	2	S37671	MHC class III hist
38	145	6.5	1872	2	S36152	MHC class III hist
39	144.5	6.4	1249	2	T14150	vesicle associated
40	144	6.4	674	2	S13301	collagen alpha 1(X
41	143	6.4	2142	2	B35098	MHC class III hist
42	142.5	6.3	1006	2	G86292	hypothetical prote
43	141.5	6.3	435	2	T15143	hypothetical prote
44	141.5	6.3	666	1	VCMVHL	env polypeptid
45	140.5	6.2	2944	2	A54849	collagen alpha 1(V

ALIGNMENTS

RESULT 1

tumor necrosis factor receptor 2-related protein - human
C:Species: Homo sapiens (man)
C>Date: 24-May-1996 #sequence-revision 24-May-1996 #text-change 17-Mar-2000
C:Accession: I54182
R:Baens, M.; Chalfant, M.; Cassiman, J.J.; Van den Berghe, H.; Marynen, P.
Genomics 16, 214-218, 1993
A:Title: Construction and evaluation of a hcdNA library of human 12p transcribed se
A:Reference number: I54182; MIMD:9325381; PMID:8486360
A:Accession: I54182
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-435 <RES>
A:Cross-references: GB:I04270; NID:g339761; PIDN:AAA36757.1; PID:g339762
C:Genetics:
A:Gene: GDB:LTBR
A:Cross-references: GDB:1230195; OMIM:600979
A:Map position: 12p13.3-12p13.1
C:Superfamily: tumor necrosis factor receptor type 1; NGF receptor repeat homology

Query Match	98.8%	Score 2220	DB 2	Length 435
Best Local Similarity	91.7%	Pred. No. 8.4e-120		
Matches 399	Conservative 0	Mismatches 0	Indels 36	Gaps 1
QY	1	MLPMWATSAPGLAMGPLVLGFLGLAASQPAVPYASENOTCRDOKEYEPOHRI	CS	60
DB	1	MLPMWATSAPGLAMGPLVLGFLGLAASQPAVPYASENOTCRDOKEYEPOHRI	CS	60
QY	61	RCPPGTIVYSAKCSRIKPTVCATCAENSNEHMYLTICQLCRPCDPVWGLEIAPCTSKR	120	
DB	61	RCPPGTIVYSAKCSRIKPTVCATCAENSNEHMYLTICQLCRPCDPVWGLEIAPCTSKR	120	
QY	121	KTCRCQPGMPCAMALECHTCELLSDCPPTBMLKDEYKGNKNCPCAGHFOYTS	180	
DB	121	KTCRCQPGMPCAMALECHTCELLSDCPPTBMLKDEYKGNKNCPCAGHFOYTS	180	
QY	181	PSARCQHTKRCENGLVEAAGTASQDPTCKNPLEPLPEKSGTMMALVLLPLAFILL	240	
DB	181	PSARCQHTKRCENGLVEAAGTASQDPTCKNPLEPLPEKSGTMMALVLLPLAFILL	240	
QY	223	-----GSLKRRPQGEENPNVAGSMEPPKAHPYPPDLVQPLLPISGD	264	
DB	241	ATVESCIMKSHPSLCRKIGSLIKRRPQGEENPNVAGSMEPPKAHPYPPDLVQPLLPISGD	300	
QY	265	VSPVSTGLPAPVLEAGVPOQSPDLTFRQLEPQSGSQYAHNTGNIHYTGSGMTTGN	324	
DB	301	VSPVSTGLPAPVLEAGVPOQSPDLTFRQLEPQSGSQYAHNTGNIHYTGSGMTTGN	360	
QY	325	IYINGVVLGGPQPGGDLPATPEPPYPIPERGDPGPGSLSTPHQEDGKAMHLATERCGA	384	
DB	361	IYINGVVLGGPQPGGDLPATPEPPYPIPERGDPGPGSLSTPHQEDGKAMHLATERCGA	420	


```

*      56484      58779: contig of 2296 bp in length
*      58780      58879: gap of unknown length
*      58880      60911: contig of 2032 bp in length
*      60912      61011: gap of unknown length
*      61012      63907: contig of 2896 bp in length
*      63908      64007: gap of unknown length
*      64008      66202: contig of 2195 bp in length
*      66203      66303: gap of unknown length
*      66303      69739: contig of 3436 bp in length
*      69739      69839: gap of unknown length
*      69839      72411: contig of 2572 bp in length
*      72411      72510: gap of unknown length
*      72510      76319: contig of 3709 bp in length
*      76319      76320: gap of unknown length
*      76320      79030: contig of 2710 bp in length
*      79030      83130: gap of unknown length
*      83130      83412: contig of 4182 bp in length
*      83412      86395: contig of 2983 bp in length
*      86395      86495: gap of unknown length
*      86495      88583: contig of 2088 bp in length
*      88583      90972: gap of unknown length
*      90972      91073: contig of 2290 bp in length
*      91073      93875: gap of unknown length
*      93875      93876: contig of 2803 bp in length
*      93876      98771: gap of unknown length
*      98771      98872: contig of 4796 bp in length
*      98872      102000: gap of unknown length
*      102000      102099: contig of 3128 bp in length
*      102099      105352: gap of unknown length
*      105352      105453: contig of 3253 bp in length
*      105453      109873: gap of unknown length
*      109873      109974: contig of 4421 bp in length
*      109974      112881: gap of unknown length
*      112881      112882: contig of 2908 bp in length
*      112882      112882: gap of unknown length
*      112882      117280: contig of 4298 bp in length
*      117280      117380: gap of unknown length
*      117380      121635: contig of 4255 bp in length
*      121635      121735: gap of unknown length
*      121735      125920: contig of 4186 bp in length
*      125920      125921: gap of unknown length
*      125921      126021: contig of 3889 bp in length
*      126021      130009: gap of unknown length
*      130009      135724: contig of 5714 bp in length
*      135724      135823: gap of unknown length

```

Query Match 24.7% Score 142; DB 2; Length 187998;
 Best Local Similarity 81.2% Pred. No. 4.5e-25;
 Matches 177; Conservative 0; Mismatches 40; Indels 1; Gaps 1;

```

QY 359 CCGGGGGGCTATGACTACTGCGACATCTACTATGACAGTGGAGGAG 418
      | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 154396 GCGACAGGCTCTGTGACTGTACCGGCAATATCTACTATGAGGAG 154455

QY 419 GACACCGGGGCTCTGAGACCTCCAGCTACCCCGAAGTCCATCCCGGAG 478
      | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 154456 GAACACGGGGGCTCTGAGACCTCCAGCTACCCCGAAGTCCCGGAG 154515

QY 479 AGGGGGGACCTGGCCCTCCGGGGCTCTACACCCCGAGAGAGTGGAGG 538
      | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 154516 AGGGAGGCGCTGGCCCTCTGAGCTGTCTACACCTTACAGAGAGTGGAG 154575

QY 539 ACCTACCGGAGACAGAG-CACGTGTGTGGCAGACCCCTC 575
      | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 154576 ATCTGGCTGAGACAGAGACTAGGTGCGAAGACCTC 154613

RESULT 13
LOCUS HUMTUMNEC 1605 bp mRNA linear PRI 14-JAN-1995
DEFINITION Homo sapiens (clone NCD18) tumor necrosis factor receptor related
            protein mRNA, complete exon and repeat region.

```

```

ACCESSION L04489
VERSION L04489.1 GI:340022
KEYWORDS tumor necrosis factor receptor related protein.
SOURCE Homo sapiens (tissue library: hncdna tc651) CDNA to mRNA.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1605)
AUTHORS Baens,M., Chalfanet,M., Cassiman,J.J., van den Berghe,H. and
            Marynen,P.
TITLE Construction and evaluation of a hncDNA library of human 12p
JOURNAL Transcribed sequences derived from a somatic cell hybrid
MEDLINE Genomics 16 (1), 214-218 (1993)
PUBMED 93252381
FEATURES
SOURCE Location/Qualifiers
            1..1605
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /map="12p13"
               /cell_line="M28 somatic cell hybrid"
               /tissue_lib="hncdna tc651"
               1..1605
               /gene="TNFR"
               /gene="TNFR"
               1..105
               /partial
               /note="putative"
               /rpt_family="Alu"
               /rpt_type="dispersed"
               250..444
               /gene="TNFR"
               /note="L04270 sequence homologue"
               /evidence="experimental"
BASE COUNT 472 a 394 c 323 g 416 t
ORIGIN
Query Match 14.4% Score 83; DB 9; Length 1605;
Best Local Similarity 100.0% Pred. No. 3.1e-10;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGTGGAGGAGCTCTCAGGACCTGCCAGACACACCTGCAAAATCATTAGAG 60
      | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 363 CTGTGGAGGAGCTCTCAGGACCTGCCAGACACACCTGCAAAATCATTAGAG 422

QY 61 CCACGTGCCCCGAGAGATGTCAG 83
      | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 423 CCACGTGCCCCGAGAGATGTCAG 445

RESULT 14
LOCUS HS27B9F/c 193 bp DNA linear PRI 19-OCT-1995
DEFINITION H.sapiens Cpg island DNA genomic MseI fragment, clone 27b9, forward
            read cpg27b9.f.tla.
ACCESSION Z60528
VERSION Z60528.1 GI:1032632
KEYWORDS Cpg island; genomic MseI fragment.
SOURCE Homo sapiens
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 193)
AUTHORS Macdonald,M., Huckle,E., Wilkinson,P. and Micklem,G.
TITLE Direct Submission
JOURNAL Submitted (16-Oct-1995) The Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1RO, England. E-mail contact: humquerry@sanger.ac.uk
REFERENCE 2 (bases 1 to 193)
AUTHORS Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
TITLE Purification of Cpg islands using a methylated DNA binding column
JOURNAL Nat. Genet. 6 (3), 236-244 (1994)
MEDLINE 94282070
PUBMED 8012384
COMMENT Vector: pGEM-5zf(-)

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE
AUTHORS

1 (bases 1 to 187998)

Mozny,D.M., Adams,C., Adio-Oduola,B., Ali-oshman,F.R., Allen,C., Alshrocks,S.L., Amaralunge,H.C., Are,J.R., Ayale,M., Banks,T., Barbieri,J., Benton,J., Bimagne,K., Blankenburg,K., Bonini,D., Bouck,J., Bowler,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Caron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthatte,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escoto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,T., Johnson,R., Jolivet,S., Joudan,S., Karlsson,E., Kelly,S., Khan,U., King,L., Koryah,J., Kovar,C., Kratochvic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozdo,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Mohabhat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,S., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokwenu,S., Oguh,M., Okunolu,G., Ocranaye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Severy,G., Scherfer,S., Scott,G., Shen,H., Shoshitari,N., Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansy,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,M., Wang,O., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wellington,S., Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G., and Gibbs,R.

TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL

Unpublished
2 (bases 1 to 187998)

Direct Submission
Submitted (02-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 187998)

REFERENCE
TITLE
JOURNAL

Worley,K.C.
Direct Submission
Submitted (23-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT
JOURNAL

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GVSA
Center clone name: CH230-100N7
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 133860 bases at least Q40
Consensus quality: 142697 bases at least Q30
Consensus quality: 148841 bases at least Q20

***** NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
NOTE: This is a 'working draft' sequence. It currently
consists of 64 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1	1031:	contig of 1031 bp in length
1032	1131:	gap of unknown length
1132	2334:	contig of 1203 bp in length
2335	2434:	gap of unknown length
2435	3488:	contig of 1054 bp in length
3489	3588:	gap of unknown length
3589	4860:	contig of 1272 bp in length
4861	4960:	gap of unknown length
4961	6021:	contig of 1061 bp in length
6022	6121:	gap of unknown length
6122	7174:	contig of 1053 bp in length
7175	7274:	gap of unknown length
7275	8532:	contig of 1258 bp in length
8533	8632:	gap of unknown length
8633	9711:	contig of 1079 bp in length
9712	9811:	gap of unknown length
9812	11308:	contig of 1497 bp in length
11309	11408:	gap of unknown length
11409	12604:	contig of 1196 bp in length
12605	12704:	gap of unknown length
12705	14114:	contig of 1410 bp in length
14115	14214:	gap of unknown length
14215	16706:	contig of 2492 bp in length
16707	16806:	gap of unknown length
16807	17875:	contig of 1069 bp in length
17876	17975:	gap of unknown length
17975	19654:	contig of 1679 bp in length
19655	19754:	gap of unknown length
19755	21683:	contig of 1929 bp in length
21684	21783:	gap of unknown length
21784	22871:	contig of 1088 bp in length
22872	22971:	gap of unknown length
22972	24666:	contig of 1695 bp in length
24667	24766:	gap of unknown length
24767	26943:	contig of 2177 bp in length
26944	27043:	gap of unknown length
27044	29057:	contig of 2014 bp in length
29058	29157:	gap of unknown length
29158	30973:	contig of 1816 bp in length
30974	31073:	gap of unknown length
31074	32838:	contig of 1765 bp in length
32839	32938:	gap of unknown length
32939	34616:	contig of 1678 bp in length
34617	37116:	gap of unknown length
37117	37198:	contig of 2482 bp in length
37199	37298:	gap of unknown length
37299	39477:	contig of 2179 bp in length
39478	39577:	gap of unknown length
39578	41815:	contig of 2238 bp in length
41816	41915:	gap of unknown length
41916	44488:	contig of 2573 bp in length
44489	44588:	gap of unknown length
44589	46062:	contig of 1474 bp in length
46063	46162:	gap of unknown length
46163	48328:	contig of 2166 bp in length
48329	48428:	gap of unknown length
48429	51395:	contig of 2967 bp in length
51396	51495:	gap of unknown length
51496	54159:	contig of 2664 bp in length
54160	54259:	gap of unknown length
54260	56383:	contig of 2124 bp in length
56384	56483:	gap of unknown length

[illegible]


```

repeat_region complement(17193..17563)
repeat_region /rpt_family="THEIA-Int"
repeat_region complement(17564..17915)
repeat_region /rpt_family="THEIA"
repeat_region complement(18210..18342)
repeat_region /rpt_family="FIAM_A"
gene join(18391..18519,18744..18902,19603..19853)
gene /gene="Vesicle-associated membrane protein 1 (Syb1)"
repeat_region 19248..19275
repeat_region /rpt_family="(CA)n"
repeat_region 19276..19351
repeat_region /rpt_family="LIME4"
repeat_region complement(19296..19475)
repeat_region /rpt_family="LIMCa"
repeat_region 19914..19933
repeat_region /rpt_family="(CCTG)n"
STS 20287..20651
STS /standard_name="GDB:214816"
STS 20290..20527
STS /standard_name="sts-M36200"
STS 20459..20650
STS /standard_name="sts-M36200"
STS 20841..21018
STS /standard_name="SHG-557"
STS 20971..21045
STS /standard_name="D12S111E"
STS 22216..22330
STS /standard_name="Cda1eb01"
STS 22518..22670
STS /standard_name="stSG4862"
gene complement(join(22595..22710,23806..23889,25797..26099,
27000..27338,31028..31297,31447..31677,32432..32495))

```

```

Query Match 41.6% Score 239 DB 9 Length 140026
Best Local Similarity 100.0% Pred. No. 2.7e+49
Matches 239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 337 GGTACCATGGCATTCATGTCACCGGGGCTCTATGACTATACGCAACATCTACATC 396
Db 94085 GGTACCATGGCATTCATGTCACCGGGGCTCTATGACTATACGCAACATCTACATC 94026
QY 397 TACATGACACGAGTACGAGGGGAGCAACCGGGTCTGAGAGACCTCCAGTACCCCGAA 456
Db 94025 TACATGACACGAGTACGAGGGGAGCAACCGGGTCTGAGAGACCTCCAGTACCCCGAA 93966
QY 457 CCTCCATACCCCATTCCTCCGAGAGGGGAGACCTCCGCGGCTCTACACCCAC 516
Db 93965 CCTCCATACCCCATTCCTCCGAGAGGGGAGACCTCCGCGGCTCTACACCCAC 93906
QY 517 CAGGAGATGAGCAGGCTTGACACCTAGAGAGAGACGACGCTGGTGACACCCCTC 575
Db 93905 CAGGAGATGAGCAGGCTTGACACCTAGAGAGAGACGACGCTGGTGACACCCCTC 93847

```

```

RESULT 11
AC128082/c 178228 bp DNA linear HTG 19-JUL-2002
LOCUS Rattus norvegicus clone CH230-362C16, *** SEQUENCING IN PROGRESS
DEFINITION *** 64 unordered pieces.
ACCESSION AC128082
VERSION AC128082.1 GI:21908679
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Rattus norvegicus.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 178228)

```

```

REFERENCE
AUTHORS Muzny/D.M., Adams/C., Adio-Oduola/B., Ali-Osman/F.R., Allen/C.,
Alshrooke/S.L., Amaralunge/H.C., Are/J.R., Ayele/M., Banks/T.,
Barbarta/J., Benton/J., Bimoge/K., Blankenburg/K., Bonnin/D.,
Bouck/J., Bowie/S., Brieva/M., Brown/E., Brown/M., Bryant/N.P.,
Buhay/C., Burch/P., Burkett/C., Burrell/K.L., Byrd/N.C.,

```

```

TITLE JOURNAL
REFERENCE JOURNAL
AUTHORS JOURNAL
TITLE JOURNAL
COMMENT

```

Carton,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dahorne,R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,R.J.,
Eamharth,C., Edgar,D., Edwards,C.C., Elhaj,C., Escoto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Licharge,O., Lien,C., Liu,J., Liu,M., Louissege,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,D.,
Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenwo,S., Ogun,M., Okunolu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rivers,M., Rojas,A., Rojoubkan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shooshari,N., Sisto,I.,
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Umanli,K., Vasquez,L., Vera,Y., Villalton,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wellington,S.,
Williams,G., Williamson,A., Wlezyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G., and Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 178228)
Worley,K.C.

Direct Submission
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: GZQW
Center clone name: CH230-362C16

----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 116771 bases at least Q40
Consensus quality: 123909 bases at least Q30
Consensus quality: 129462 bases at least Q20

----- NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
NOTE: This is a 'working draft' sequence. It currently
consists of 64 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

```

1 1173: contig of 1173 bp in length
* 1174 1273: gap of unknown length
* 1274 2435: contig of 1162 bp in length
* 2436 2535: gap of unknown length

```


/translation="MRLPRASSPCGLAMGBLLGLSLVASOPOLVPPYRIENQTCM
DQREYEEPMHIVCCSCRCPEGEFVAVCSRSODTVCKTCPEHNSYNEHMHSLSTCOLCR
PCDIVLGEFEVAPCTSDRAKRCOCFQMGSCVYLIDNEVCHEERLYLCOPTAEAVTD
EIMDTVNCVPCPKRGHONTSPPRARCOPHTRCIOGLVEAARGTSYSDTCKNPPRP
GAMLLAIIILSLYLFTTIVTLACAMMRHSILCRKIGTLKLRHPEGEESPCCAPRAD
PHRPDLAEPLPMSGDLSPSPAGPTAPASLEEVYLQOOSPLVQARELEAPGEGHGYA
HGANGIHVTVGVSIVTGNITLYNGPVLGCTRGGRDPAPPEPPYPTPEEGAPGSPSELS
TPYQEDGKAMHLEETETLIGCDL"

BASE COUNT 316 a 537 c 455 g 306 t

ORIGIN

Query Match 48.5%; Score 278.8; DB 10; Length 1614;
Best Local Similarity 80.0%; Pred. No. 3.9e-59;
Matches 328; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 146 CTCGGAAGGCCCATCATCTTCCCTGACTGCTAGACAGCCCATCTCCATCCCTTCTGGAG 205
DB 1000 CTCGGAAGGCCCATCATCTTCCCTGACTGCTAGACAGCCCATCTCCATCTGAG 1059
QY 206 ATGTTTCCCAATATCCACTGCGGCTCCCGCAGCCCATGTTTGGAGGAGGCTGCCGC 265
DB 1060 ACTGTTCCCAATATCCACTGCGGCTCCCGCAGCCCATGTTTGGAGGAGGCTGCCGC 1119
QY 266 AACAGCAGAGTCTCTGACCTGACAGAGGAGCCGCAATTTGGAACCCGGGAGACAGACC 325
DB 1120 AACAGCAGAGTCTCTGACCTGACAGAGGAGCCGCAATTTGGAACCCGGGAGACATGCC 1179
QY 326 AGGTGGCCCAAGTACCATGATGATGACAGGAGGAGGCTGATGATGATGATGATGATG 385
DB 1180 AGGTGGCCCAAGTACCATGATGATGACAGGAGGAGGCTGATGATGATGATGATGATG 1239
QY 386 ACATCTACATATACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 445
DB 1240 ACATCTACATATACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1299
QY 446 CTACCCCGGAACCTTCATACCTCCGGAAGAGGGGAGCCCTGCGCTCCCGGGCTCT 505
DB 1300 CTCCCGCGGAGCTTCATACCTCCGGAAGAGGGGAGCCCTGCGCTCTGAGCTGT 1359
QY 506 CTACACCCCGGAGGAGATGAGCAAGGCTTGACCTGAGGAGGAGAGAGAG 555
DB 1360 CTACACCCCGGAGGAGATGAGCAAGGCTTGACCTGAGGAGGAGAGAGAG 1409

RESULT 9
MMU29173 2076 bp mRNA linear ROD 28-JUN-1995
DEFINITION Mus musculus lymphotoxin-beta receptor mRNA, complete cds.
ACCESSION U29173
VERSION U29173.1 GI:881620
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2076)
AUTHORS Force,W.R., Williams-Abbott,L., Browning,J., Hession,C., Tizard,R.
and Ware,C.F.
TITLE Cloning and Characterization of the Mouse Lymphotoxin-beta Receptor
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 2076)
TITLE Direct Submission
JOURNML Submitted (15-JUN-1995) Walker R. Force, Biomedical Sciences, U.C.
Riverside, Riverside, CA 92521-0121, USA
FEATURES
Location/Qualifiers
1..2076
/organism="Mus musculus"
/strain="CWB"
/db_xref="taxon:10090"
/sex="female"
/tissue_type="lung"
170..1417

CDS

/note="transmembrane protein"
/codon_start=1
/product="lymphotoxin-beta receptor"
/protein_id="AA68964.1"
/db_xref="GI:881621"

/translation="MRLPRASSPCGLAMGBLLGLSLVASOPOLVPPYRIENQTCM
DQREYEEPMHIVCCSCRCPEGEFVAVCSRSODTVCKTCPEHNSYNEHMHSLSTCOLCR
PCDIVLGEFEVAPCTSDRAKRCOCFQMGSCVYLIDNEVCHEERLYLCOPTAEAVTD
EIMDTVNCVPCPKRGHONTSPPRARCOPHTRCIOGLVEAARGTSYSDTCKNPPRP
GAMLLAIIILSLYLFTTIVTLACAMMRHSILCRKIGTLKLRHPEGEESPCCAPRAD
PHRPDLAEPLPMSGDLSPSPAGPTAPASLEEVYLQOOSPLVQARELEAPGEGHGYA
HGANGIHVTVGVSIVTGNITLYNGPVLGCTRGGRDPAPPEPPYPTPEEGAPGSPSELS
TPYQEDGKAMHLEETETLIGCDL"

BASE COUNT 436 a 652 c 597 g 391 t

ORIGIN

Query Match 48.5%; Score 278.8; DB 10; Length 2076;
Best Local Similarity 80.0%; Pred. No. 3.9e-59;
Matches 328; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 146 CTCGGAAGGCCCATCATCTTCCCTGACTGCTAGACAGCCCATCTCCATCCCTTCTGGAG 205
DB 984 CTCGGAAGGCCCATCATCTTCCCTGACTGCTAGACAGCCCATCTCCATCTGAG 1043
QY 206 ATGTTTCCCAATATCCACTGCGGCTCCCGCAGCCCATGTTTGGAGGAGGCTGCCGC 265
DB 1044 ACTGTTCCCAATATCCACTGCGGCTCCCGCAGCCCATGTTTGGAGGAGGCTGCCGC 1103
QY 266 AACAGCAGAGTCTCTGACCTGACAGAGGAGCCGCAATTTGGAACCCGGGAGACAGACC 325
DB 1104 AACAGCAGAGTCTCTGACCTGACAGAGGAGCCGCAATTTGGAACCCGGGAGACATGCC 1163
QY 326 AGGTGGCCCAAGTACCATGATGATGACAGGAGGAGGCTGATGATGATGATGATGATG 385
DB 1164 AGGTGGCCCAAGTACCATGATGATGACAGGAGGAGGCTGATGATGATGATGATGATG 1223
QY 386 ACATCTACATATACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 445
DB 1224 ACATCTACATATACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1283
QY 446 CTACCCCGGAACCTTCATACCTCCGGAAGAGGGGAGCCCTGCGCTCCCGGGCTCT 505
DB 1284 CTCCCGCGGAGCTTCATACCTCCGGAAGAGGGGAGCCCTGCGCTCTGAGCTGT 1343
QY 506 CTACACCCCGGAGGAGATGAGCAAGGCTTGACCTGAGGAGGAGAGAGAG 555
DB 1344 CTACACCCCGGAGGAGATGAGCAAGGCTTGACCTGAGGAGGAGAGAGAG 1393

RESULT 10
AC005840/c 140026 bp DNA linear PRI 20-OCT-2000
DEFINITION Homo sapiens complete sequence of a PAC clone RPI-102E24 containing
SYB1, CD27, and SCNN1A genes.
ACCESSION AC005840
VERSION AC005840.2 GI:10938025
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 140026)
AUTHORS Montgomery,K.T., Lau,S.T. and Kucherlapati,R.
TITLE High throughput Sequencing of Human Chromosome 12
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 140026)
TITLE Direct Submission
JOURNML Submitted (22-OCT-1998) Department of Molecular Biology, Albert
Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY
10461, USA
REFERENCE 3 (bases 1 to 140026)
AUTHORS Montgomery,K.T., Lau,S.T. and Kucherlapati,R.

DB	Accession	Source	Organism	Reference	Authors	Journal	Remark	Comment
Db	1326	CCAGCTACCCCCGAACTCCATACCCATTCGCCGAAAGAGGGGAGACCCCTGCCGG	1385					
Oy	502	CTCTCTACACCCACACGAGAGATGCGACGGCTTGACCTACCGAGACAGACACTGT	561					
Db	1386	CTCTCTACACCCACACGAGAGATGCGACGGCTTGACCTACCGAGACAGACACTGT	1445					
Oy	562	GGTGGCACACCCCTC	575					
Db	1446	GGTGGCACACCCCTC	1459					
RESULT 7								
LOCUS	BC026262	2161 bp	mRNA	linear	PRI 08-APR-2002			
DEFINITION	Homo sapiens, lymphotoxin beta receptor (TNFR superfamily, member							
ACCESSION	BC026262							
VERSION	BC026262.1							
KEYWORDS	MG.							
SOURCE	Homo sapiens.							
ORGANISM	Homo sapiens.							
REFERENCE	1 (bases 1 to 2161)							
AUTHORS	Strausberg, R.							
TITLE	Direct Submission							
JOURNAL	submitted (02-APR-2002) National Institutes of Health, Mammalian							
	Gene Collection (MGC), Cancer Genomics Office, National Cancer							
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,							
	USA							
	NIH-MGC Project URL: http://mgc.ncl.nih.gov							
	Contact: MGC help desk							
	Email: cga@bbs-remail.nih.gov							
	Tissue Procurement: CLONTECH							
	CDNA Library Preparation: CLONTECH Laboratories, Inc.							
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)							
	DNA Sequencing by: Sequencing Group at the Stanford Human Genome							
	Center, Stanford University School of Medicine, Stanford, CA 94305							
	Web site: http://www-shgc.stanford.edu							
	Contact: (Dickson, Mark) mdcpaxil.stanford.edu							
	Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,							
	R. M.							
FEATURES								
source	1. 2161							
	/organism="Homo sapiens"							
	/db_xref="locusID:4055"							
	/db_xref="taxon:9606"							
	/clone="MGC:22593 IMAGE:4703362"							
	/tissue_type="Lung"							
	/clone_lib="NIH_MGC_77"							
	/lab_host="DH10B"							
	/note="Vector: pDNR-LIB"							
	199. 1506							
	/codon_start=1							
	/product="lymphotoxin beta receptor (TNFR superfamily,							
	member 3)"							
	/protein_id="AAH26262.1"							
	/db_xref="GI:20072213"							
	/translation="MLPMTASAPGLAMGPVILGELGLAASOPAVVPYASENQTCL							
	DKREYTERPDHRIICRCRCRPGTYVSAKCSRIDTYVATCAENSYNHMYLITCOLC							
	PCDPVAGLELITAPCTSRKTRKQRCQGMFCAMALECTCELLSDPGTEALMDE							
	GKNNHCPCVCKAGHFOVTSPPSARCGPHRCENOGLEAARPGTASDITCKNPLPIL							
	PENSGTLMALVLLPLAFLLATVFSCTWKSHPSLCRKLGLSLKRRPGDEENPYAA							
	PSGFAHYPFDLVDPLPIISGDSVPSVGLPFAFVILAGVAGQDSPLDITREPL							
	PGDSQVAGHTNGIHVGTGSLITGNMIIYINGVILGSPGPGDLPATPPPIPIPEE							
	DPEPGISTPHQEDGKAMHIALETGHEGATPSSNNGPSPNOFTIDH"							

Query Match	85.9%	Score 494;	DB 9;	length 2161;
Best Local Similarity	100.0%	Pred. No. 6.9e-113;		
Matches 494;	Conservative	0;	Mismatches 0;	Indels 0;
			Gaps	0;
82	GGATGCTGCTCAGAGAGCGCTCCGACAGAGAGAGGAGCCCAATCTCTGATGCTGGAACTGG	141		
Db	973	GGATGCTGCTCAGAGAGCGCTCCGACAGAGAGAGGAGCCCAATCTCTGATGCTGGAACTGG	1032	
Qy	142	GAGCCTCCGAAGGCCCATTCATCTTCCCTGACTTGGACAGGCACATGCAATTCAT	201	
Db	1033	GAGCCTCCGAAGGCCCATTCATCTTCCCTGACTTGGACAGGCACATGCAATTCAT	1092	
Qy	202	GGAGATGTTTCCCGAGTATCCACTGGGCTCCCGACGCCAGTCTTGGAGCAGGGGTG	261	
Db	1093	GGAGATGTTTCCCGAGTATCCACTGGGCTCCCGACGCCAGTCTTGGAGCAGGGGTG	1152	
Qy	262	CGGCAACAGCAGAGTCTCTGGACCTGCACAGGAGACCCGACGTTGGAAACCCGGGAGACAG	321	
Db	1153	CGGCAACAGCAGAGTCTCTGGACCTGCACAGGAGACCCGACGTTGGAAACCCGGGAGACAG	1212	
Qy	322	AGCCAGTGGCCACGGTACCAATGGCATTCATGTCACCGCGGGTCTGATCATATCACT	381	
Db	1213	AGCCAGTGGCCACGGTACCAATGGCATTCATGTCACCGCGGGTCTGATCATATCACT	1272	
Qy	382	GGCAATATATCATATGACATGACAGTACGAGGAGGAGACACCGGGTCTGAGAACCTC	441	
Db	1273	GGCAATATATCATATGACATGACAGTACGAGGAGGAGACACCGGGTCTGAGAACCTC	1332	
Qy	442	CCAGCTACCCCGGAACCTCCATACCCCATTTCCGAAAGAGGGGAGCCCTGCGCTCCCGG	501	
Db	1333	CCAGCTACCCCGGAACCTCCATACCCCATTTCCGAAAGAGGGGAGCCCTGCGCTCCCGG	1392	
Qy	502	CTCTTACACCCACACAGAGATGGCAAGGCTTGGACACTGACGGAGACAGACACTGT	561	
Db	1393	CTCTTACACCCACACAGAGATGGCAAGGCTTGGACACTGACGGAGACAGACACTGT	1452	
Qy	562	GGTGGCAGACCCCTC	575	
Db	1453	GGTGGCAGACCCCTC	1466	

```

RESULT 8
MUSLYMHOB
LOCUS      1614 bp      mRNA      linear      ROD 24-MAY-1996
DEFINITION Mus musculus lymphotoxin-beta receptor gene, complete cds.
ACCESSION L38423
VERSION   L38423.1
KEYWORDS  lymphotoxin-beta receptor; transmembrane protein.
SOURCE    Mus musculus.
ORGANISM  Mus musculus.
REFERENCE Eukariyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus.
           1 (bases 1 to 1614)
           Nakamura, T., Tashiro, K., Nazarea, M., Nakano, T., Sasayama, S. and
           Honjo, T.
           The murine lymphotoxin-beta receptor cDNA: isolation by the signal
           sequence trap and chromosomal mapping
           Genomics 30 (2), 312-319 (1995)
           8586432
FEATURES
Source
     1..1614
        /organism="Mus musculus"
        /db_xref="taxon:10090"
        /cell_line="ST-2"
     186..1433
        /note="putative"
        /codon_start=1
        /product="lymphotoxin-beta receptor"
        /protein_id="AAB00846.1"
        /db_xref="GI:600223"

```

FEATURES	Source	Location/Qualifiers
PUBMED 8486360		1. 2136
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/tissue_type="liver"
		/tissue_lib="liver CDNA of P.M."
		169..1476
		/note="putative"
		/codon_start=1
		/product="tumor necrosis factor receptor 2 related protein"
		/protein_id="AA36757.1"
		/db_xref="GI:339762"
		/translation="MLPFAATSAAGLAWGLVGLFEFLAASQPAVPVASENOTR DDEKVEIPEPHRITCCSCRPPIGVTSACSTIRDTVCATCENSLNEWMNLITTCOLR PCDVGWGLEIEIAPCTSKRKTYQCRQPMFCFAAVALICTHCELLSDPGTEAELEKDEV GKGNHNCPCPKAGHFTQMTSSPSARCOPTHTCENOGLVEAAFGTAQSDTCKNLEPLD PEMSGTMLMAVLPLFLFLFLFTSIVSFSCIKMSHSLCRKLSILKRDPQEGPVAVG SWEPKSAHYEPDLPVLPFLISGDSVSPSGLPAAPVLEAGVPPQDSPLDTRPOLE PESROVAHGTNGIHVYGLSMETITGNLYITINGPLVGGPQPGDLPARPEPPIPIPEG DPRPGLSTRPHODGKAMHLAEMPHCATSTSNKPNQFIITHD"
BASE COUNT	446 a	706 c 608 g 376 t
ORIGIN		
Query Match	85.9%; Score 494; DB 9; Length 2136;	
Best Local Similarity	100.0%; Pred. No. 6.9e-113;	
Matches 494; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
OY	82	GGATGCTGCTCAAGAGCGCTCCGAGGGAAGAGGACCCAAATCTGTAGCTGGAACTGG 141
Db	943	GGATCGCTGCTCAAGAGCGCTCCGAGGGAAGAGGACCCAAATCTGTAGCTGGAACTGG 1002
OY	142	GAGCCTCCGAAGGCCATCATCTCTCCCTGACTGTGATACGCCACCTGTACCATTTCT 201
Db	1003	GAGCCTCCGAAGGCCATCATCTCTCCCTGACTGTGATACGCCACCTGTACCATTTCT 1062
OY	202	GGAGATGTTTCCCGAGTATTCACATGGGCTCCCGAGGCCCCAGATTTTGGAGGCAAGGGGTG 261
Db	1063	GGAGATGTTTCCCGAGTATTCACATGGGCTCCCGAGGCCCCAGATTTTGGAGGCAAGGGGTG 1122
OY	262	CCGCAACGACAGAGTCCCTGGACCTGACACAGGAGCCGACAGTTGGAAACCCGGGGAGCAG 321
Db	1123	CCGCAACGACAGAGTCCCTGGACCTGACACAGGAGCCGACAGTTGGAAACCCGGGGAGCAG 1182
OY	322	AGCCAGTGGGCCAAGGTATACCAATGGCAATGATATGACACGGCGGGTCTATGACTATTA 381
Db	1183	AGCCAGTGGGCCAAGGTATACCAATGGCAATGATATGACACGGCGGGTCTATGACTATTA 1242
OY	382	GGCAACATCTATCTATCAATATGAGACCACTACTGGGGGAGACACCGGGTCTTGGAGACCTC 441
Db	1243	GGCAACATCTATCTATCAATATGAGACCACTACTGGGGGAGACACCGGGTCTTGGAGACCTC 1302
OY	442	CCAGCTACCCCGGAACCTCCATACCCCAATTCGCCGAAGAGGGGAGACCTGGGCCCTCCGGG 501
Db	1303	CCAGCTACCCCGGAACCTCCATACCCCAATTCGCCGAAGAGGGGAGACCTGGGCCCTCCGGG 1362
OY	502	CTCTTACACCCACACAGAGAGATGGAAGAGCTTGGACCTAGCGGAGACAGAGACATGT 561
Db	1363	CTCTTACACCCACACAGAGAGATGGAAGAGCTTGGACCTAGCGGAGACAGAGACATGT 1422
OY	562	GGTGGCACACCTC 575
Db	1423	GGTGGCACACCTC 1436
RESULT 6		
LOCUS	AK027080	2148 bp mRNA linear PRI 29-SEP-2000
DEFINITION	Homo sapiens cDNA: FLJ23427 fls, clone HRC04788, highly similar to	
	HUMTNFRP Homo sapiens tumor necrosis factor receptor 2 related	
	protein mRNA.	
ACCESSION	AK027080	

VERSION	AK027080.1	GI:10440111
KEYWORDS	Oligo capping; f1s (full insert sequence).	
SOURCE	Homo sapiens primary human renal epithelial cells cDNA to mRNA, clone_11b:HRC clone:HRC04788.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
AUTHORS	Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shihabara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.	
TITLE	NEO human cDNA sequencing project	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 2148)	
AUTHORS	Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shihabara,T., Tanaka,T. and Nakamura,Y.	
TITLE	Direct Submission	
JOURNAL	Submitted (23-AUG-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@hims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)	
COMMENT	NEO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- 6 3'-end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).	
FEATURES	Location/Qualifiers	
source	1..2148	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/clone="HRC04788"	
	/cell_type="primary human renal epithelial cells"	
	/clone_11b="HRC"	
	/note="Cloning vector pME18SFT3"	
misc_feature	1..2148	
	/note="highly similar to HUNTNRFP Homo sapiens tumor necrosis factor receptor 2 related protein mRNA"	
BASE COUNT	462 a 705 c 606 g 375 t	
ORIGIN		
Query Match	85.9%; Score 494; DB 9; Length 2148;	
Best Local Similarity	100.0%; Prid. No. 6.9e-113;	
Matches 494; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	82 GGATGGCGTCTCAAGAGCGTCCGAGGAGAGGAGGACCAATCCGTGACGTGAGAGCTGG 141	
Db	966 GGATGGCTGTCTAAAGAGGGTCCGAGGAGGAGGAGCCCAATCCGTGACGTGAGAGCTGG 1025	
QY	142 GAGCCTCGAAGGGCCATATCATCTTCCCTGACTTGATGATACAGCCACTGTACCATTTCT 201	
Db	1026 GAGCCTCGAAGGGCCATATCATCTTCCCTGACTTGATGATACAGCCACTGTACCATTTCT 1085	
QY	202 GGAGATGTTTCCCGAGTATCCACTGGGCTCCCGCAGCCCGAGTTTGGAGGCAGGGGTG 261	
Db	1086 GGAGATGTTTCCCGAGTATCCACTGGGCTCCCGCAGCCCGAGTTTGGAGGCAGGGGTG 1145	
QY	262 CCGCAACAGCAGAGTCCTGTGAGCTTGACACAGGAGCGCAGTTTGGAAACCGGGAGAGC 321	
Db	1146 CCGCAACAGCAGAGTCCTGTGAGCTTGACACAGGAGCGCAGTTTGGAAACCGGGAGAGC 1205	
QY	322 AGCCAGTGGCCCGACGATACCAATGGCATTCGTACCGGGCGGTCTATGACTATCACT 381	
Db	1206 AGCCAGTGGCCCGACGATACCAATGGCATTCGTACCGGGCGGTCTATGACTATCACT 1265	
QY	382 GGCACATCTCATATTAATAATGACACAGTACGTGGGGGACACACGGGTCCTGAGAACCTC 441	
Db	1266 GGCACATCTCATATTAATAATGACACAGTACGTGGGGGACACACGGGTCCTGAGAACCTC 1325	
QY	442 CCAGGTACCCCGGACACTTCATACCCCATTCCTCCGGAAGAGGGGAGCCCTGAGCCTCCGCG 501	

TITLE	NEDO human cDNA sequencing project			
JOURNAL	unpublished			
REFERENCE	2 (bases 1 to 2091)			
AUTHORS	Iisogi,T. and Yamamoto,J.			
TITLE	Direct Submission			
JOURNAL	Submitted (04-JUL-2002) Takao Iisogi, FLJ Project(HRI Team); 2-6-7			
COMMENT	Kazuo-Kametani, Kisanazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.			
FEATURES	Location/Qualifiers			
source	1..2091.			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/clone="D3OST1000238"			
	/cell_type="CD34+ Cells"			
	/clone_1lb="D3OST1"			
	/note="Cloning vector: pME18SFL3-mRNA from CD34+ cells after 3-days ODF induction."primary culture, CD34+ Cells"			
BASE COUNT	440 a 691 c 594 g 366 t			
ORIGIN				
Query Match	85.9%	Score 494;	DB 9;	Length 2091;
Best Local Similarity	100.0%	Pred. No. 7e-113;		
Matches 494;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	82	GGATGCGTCTCAAGAGCGTCCGAGGAGAGAGACCAATCTGTAGCTTGAAGCTGG	141	
Db	930	GGATGCGTCTCAAGAGCGTCCGAGGAGAGAGACCAATCTGTAGCTTGAAGCTGG	989	
QY	142	GAGCTCCGAGAGGCCATTCATCTTCCCTGACTTGTGTAAGCACTGCTACCACTTCT	201	
Db	990	GAGCTCCGAGAGGCCATTCATCTTCCCTGACTTGTGTAAGCACTGCTACCACTTCT	1049	
QY	202	GGAGATGTTTCCCGCATATCCATGCTGGGTCCCGGAGCCGAGTTTGGAGGACAGGGGTG	261	
Db	1050	GGAGATGTTTCCCGCATATCCATGCTGGGTCCCGGAGCCGAGTTTGGAGGACAGGGGTG	1109	
QY	262	CCGCAACAGCAGAGTCTCTGTGACCTGACAGGAGACCGCACTTGGAAACCCGGGAGCAG	321	
Db	1110	CCGCAACAGCAGAGTCTCTGTGACCTGACAGGAGACCGCACTTGGAAACCCGGGAGCAG	1169	
QY	322	AGCCAGGAGGCCACAGGTAACAATGGCAATTCATGACACCGGGGTCTATGACTATCACT	381	
Db	1170	AGCCAGGAGGCCACAGGTAACAATGGCAATTCATGACACCGGGGTCTATGACTATCACT	1229	
QY	382	GGCAACATCTAATCTACATGAGACCACTAGCGGGGAGCACACCGGGTCTTGGAAACCTC	441	
Db	1230	GGCAACATCTAATCTACATGAGACCACTAGCGGGGAGCACACCGGGTCTTGGAAACCTC	1289	
QY	442	CCAGTACACCCCGAAGCTTCATACCCCATTCGCCGAAGAGGGGAGACCCCTGCTCCCGGG	501	
Db	1290	CCAGTACACCCCGAAGCTTCATACCCCATTCGCCGAAGAGGGGAGACCCCTGCTCCCGGG	1349	
QY	502	CTCTTACACCCCGAAGAGATGGCAAGGTTTGGACCTTAGCGAGACAGACACTGT	561	
Db	1350	CTCTTACACCCCGAAGAGATGGCAAGGTTTGGACCTTAGCGAGACAGACACTGT	1409	
QY	562	GGTGCACACCCCTC 575		

[illegible]

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:04:15 ; Search time 1072.25 seconds
(without alignments)
15606.600 Million cell updates/sec

Title: US-09-917-372-2_COPY_804_1378

Perfect score: 575
Sequence: 1 ctggtgagcagctccagg.....cactgtgtgtccacaccctc 575

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 205640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pin:*
35: em_htg_rnd:*
36: em_htg_mam:*
37: em_htg_vit:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	494	85.9	2091	9	AK095418	AK095418 Homo sapi
2	494	85.9	2136	6	AX331906	AX331906 Sequence
3	494	85.9	2136	6	AX332212	AX332212 Sequence
4	494	85.9	2136	6	AX409488	AX409488 Sequence
5	494	85.9	2136	6	AX409488	L04270 Homo sapien
6	494	85.9	2136	9	HUMTNRPR	AK027080 Homo sapi
7	494	85.9	2161	9	BC026262	BC026262 Homo sapi
8	278.8	48.5	1614	10	MUSLYMPHOB	L38423 Mus musculu
9	278.8	48.5	2076	10	AC005840	U29173 Mus musculu
10	239	41.6	140026	9	AC005840	AC005840 Homo sapi
11	161	28.0	178228	2	AC128082	AC128082 Rattus no
12	142	24.7	187998	2	AC125909	AC125909 Rattus no
13	83	14.4	1605	9	HUMTUNMEC	L04489 Homo sapien
14	77.4	13.5	193	9	HS2789P	Z60528 H.sapiens C
15	42.4	7.4	125020	9	AF429315	AF429315 Homo sapi
16	42.2	7.3	2048	10	MMU030798	U30798 Mus musculu
17	40	7.0	125020	9	AF429315	AF429315 Homo sapi
18	39	6.8	61851	9	AC022240	AC022240 Homo sapi
19	39	6.8	179357	9	AC074191	AC074191 Homo sapi
20	39	6.8	187998	2	AC125909	AC125909 Rattus no
21	38.8	6.7	49972	2	AC100021	AC100021 Mus muscu
22	38.8	6.7	192783	2	AC113951	AC113951 Mus muscu
23	38.6	6.7	96397	2	AC125957	AC125957 Rattus no
24	37.8	6.6	6349	5	GCTNY	X99062 G.gallus mr
25	37.8	6.6	159950	2	AC016413	AC016413 Homo sapi
26	37.8	6.6	191426	2	AC026086	AC026086 Homo sapi
27	37.6	6.5	112311	9	AL358780	AL358780 Human DNA
28	37.4	6.5	66808	1	SAU421825	AJ421825 Stigmatel
29	37.4	6.5	185464	2	AC129237	AC129237 Rattus no
30	37.2	6.5	42377	2	AC008375	AC008375 Homo sapi
31	37.2	6.5	127811	2	AC008375	AC009278 Homo sapi
32	37.2	6.5	171483	2	AC009278	AC009194 Homo sapi
33	37.2	6.5	173669	2	AC017110	AC008733 Homo sapi
34	37.2	6.5	179203	9	AC008733	AC073640 Homo sapi
35	37.2	6.5	189486	9	AC008733	AC104561 Homo sapi
36	37.2	6.5	201981	2	AC073640	AP000900 Homo sapi
37	37	6.4	155366	2	AC104561	AL845261 Mus muscu
38	37	6.4	178022	2	AP000900	AC106606 Rattus no
39	36.8	6.4	184545	2	AL845261	AC118809 Rattus no
40	36.6	6.4	151437	2	AC118809	AC022960 Homo sapi
41	36.6	6.4	171051	9	AX002166	AX002166 Sequence
42	36.4	6.3	333	6	HSY12721	Y12721 H.sapiens m
43	36.4	6.3	333	9	BD012343	BD012343 Genes rel
44	36.4	6.3	2497	6	BD012343	
45	36.4	6.3	2497	6	BD012343	

ALIGNMENTS

RESULT 1
AK095418 2091 bp mRNA linear PRI 15-JUL-2002
LOCUS AK095418 Homo sapiens CDNA FLJ18099 fls, clone D3OST1000238, highly similar
DEFINITION to LYMPHOXIN-BETA RECEPTOR PRECURSOR.
ACCESSION AK095418
VERSION AK095418.1 GI:21754669
KEYWORDS oligo capping; fls (full insert sequence).
SOURCE Homo sapiens CD34+ Cells CDNA to mRNA, Clone Jlib:D3OST1
clone:D3OST1000238.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 Oshima,A., Takahashi-Fuji,I., Tanase,T., Imose,N., Takeuchi,K.,
AUTHORS

[illegible]

```

RESULT 14
US-09-907-372-13
: Sequence 13, Application US/09907372
: Patent No. US20020068242A1
: GENERAL INFORMATION:
: APPLICANT: Lal, Preeti G.
: APPLICANT: Warren, Bridget A.
: TITLE OF INVENTION: THE RECEPTOR 2 RELATED PROTEIN VARIANT
: FILE REFERENCE: PC-0050 US
: CURRENT APPLICATION NUMBER: US/09/907,372
: CURRENT FILING DATE: 2001-07-27
: NUMBER OF SEQ ID NOS: 20
: SOFTWARE: PERL Program
: SEQ ID NO 13
: LENGTH: 206
: type: DNA
: ORGANISM: Rattus norvegicus
: FEATURE:
: NAME/KEY: misc feature
: OTHER INFORMATION: Incyte ID No. US20020068242A1 700302531H1
US-09-907-372-13

```

Query Match	20.9%;	Score 135.8;	DB 10;	Length 206;
Best Local Similarity	81.0%;	Pred. No. 1.1e-32;		
Matches 158; Conservative	0;	Mismatches 37;	Indels 0;	Gaps 0;

QY	455	CTGGACCTGACACAGGAGACCCGACTTTGGAAACCCGGGGAGCGAGACCAAGGTGGCCACAGGT	514
Db	2	CTGATTCAGAGCCAGGGAGACTGGAGGCTGAGGCTGAGGCACTGAGCCAGGTGGCCACAGGT	61
QY	515	ACCAANTGGACATTCATGTCACCGCGGGGTCTATGACTATCACTGGCAACATCTACATCTAC	574
Db	62	GCGAATGGCATTTCACGTGAGCCGAGGCGCTGTGTACTCTCAACCGGCATATCTCATATATAC	122
QY	575	AATGGACCACTACTGGGGGGGACCAACCGGGTCTTGAGACCTCCAGCTACCCCCGAACCT	634
Db	122	AATGGGCGCAGTGTGTGGGGGAGAACCGGGGCGCTTGAAACCTCTCAGCTCCCCCTGAGACCT	181
QY	635	CCATACCCGACTGCC	649
Db	182	CCATACCCGACTGCC	196

```

RESULT 15
US-09-907-372-15
: Sequence 15, Application US/09907372
: Patent No. US20020068242A1
:
: GENERAL INFORMATION:
: APPLICANT: Lal, Preeti G.
: APPLICANT: Warren, Bridget A.
: TITLE OF INVENTION: THE RECEPTOR 2 RELATED PROTEIN VARIANT
: FILE REFERENCE: PC-0050 US
: CURRENT APPLICATION NUMBER: US/09/907,372
: CURRENT FILING DATE: 2001-07-27
: NUMBER OF SEQ ID NOS: 20
:
: SOFTWARE: PERL Program
: SEQ ID NO 15
:
: LENGTH: 471
: TYPE: DNA
:
: ORGANISM: Rattus norvegicus
:
: FEATURE:
: NAME/KEY: misc.feature
:
: OTHER INFORMATION: Incyte ID No. US20020068242A1 702022948H1
: US-09-907-372-15

```

Query Match	20.9%;	Score 135.8;	DB 10;	Length 471;
Best Local Similarity	-76.3%;	Pred. No. 1.3e-32;		
Matches 167; Conservative	0;	Mismatches 52;	Indels 0;	Gaps 0;

OY	17	GAGGTACTCTTGATGACGGCCCGGCACATTAACCCAGCTCAAAATGAATGGGGAAG	76
Db	205	GAGGGGCTTGTACTCTGTACGGCCCTGGGCACAGAAGCTGAGGTACACAGATTGAATTTATGAT	264
OY	77	GGTAACAACAACACTGTCGCCCTGCGAAGGACAGGGCAGTTCAGAAATPACTCTCTCCCGAGC	138
Db	265	ACTGAAGTACACTGTGTGTCCTCTGTAAGGCACAGACATTTCCAGAACACGTCTCTCCCCAGA	324
OY	137	GCCCGCTGCCAGGCCCCACACCAGGTGTGTAGAACCAAGTCTGGTGGAGGACGCTCCAGGC	196
Db	325	GCCCGCTGTCAACCCCAACACAGGTGTGTAGAGCCACAGGGCCTGGGTGGAGGACAGCTTCAGGT	384
OY	197	ACTGCCAGTCCGACACACCAACCTGGAAAAATCATTTAGAG	235
Db	385	ACCTGTACTCTGACACACATTTGTAAAATAATCCACCGAG	423

Search completed: April 15, 2003, 18:45:43
Job time : 52.2209 secs

```
QY 1 CGAGTGACACTGCGAGCTACTTCTGACTGCGCGCTGCGACTGAAGCCGAGCTCAA 60
    |||
Db 239 CGAGTGACACTGCGAGCTACTTCTGACTGCGCGCTGCGACTGAAGCCGAGCTCAA 238
QY 61 AGATGAAGTTGGGAAGGTTAAACAACCACTGCTGCTCCCTGCAAGGCGAGCTT--CCAG 118
    |||
Db 299 AGATGAAGTTGGGAAGGTTAAACAACCACTGCTGCTCCCTGCAAGGCGAGCTTCCAGA 358
QY 119 AATACCTCCTCCCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCG 178
    |||
Db 359 AGTACCTCCTCCCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCG 418
QY 179 GTGAGGCGAGCTCCTGAGGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCG 238
    |||
Db 419 GTGAGGCGAGCTCCTGAGGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCG 478
QY 239 CTG 241
    |||
Db 479 CTG 481
```

```
RESULT 11
US-09-907-372-16
; Sequence 16, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Warren, Bridget A.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 16
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 702245091H1
US-09-907-372-16
```

```
Query Match 26.5%; Score 171.8; DB 10; Length 371;
Best Local Similarity 81.0%; Pred. No. 8.2e-44;
Matches 200; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
```

```
QY 402 GGCCTCCCGCAGCCCGAGTTTGGAGGCGAGGGTCCGCAACAGAGTCTCTGAGCC 461
    |||
Db 4 GATCCCGCAGCCCGCAGTTTGGAGGCGAGGGTCCGCAACAGAGTCTCTGAGCC 63
QY 462 TGACCAAGGAGCGGCACTTGAACCGGCGGAGCAAGAGCCAGGTTGCCACGGTACCAATG 521
    |||
Db 64 AGCCCAAGGAGCGGCACTTGAACCGGCGGAGCAAGAGCCAGGTTGCCACGGTACCAATG 123
QY 522 GCATTCATGTCACCGGCGGCTATGACTATGCTATGCTGCAACATCTAATCTACAAATGGAC 581
    |||
Db 124 GCATTCATGTCACCGGCGGCTATGACTATGCTATGCTGCAACATCTAATCTACAAATGGAC 183
QY 582 CAGTACTGGGGGGGAGCAACCGGCTCTGAGAGACCTCCAGCTACCCCGAACTCTCAATACC 641
    |||
Db 184 CAGTACTGGGGGGGAGCAACCGGCTCTGAGAGACCTCCAGCTACCCCGAACTCTCAATACC 243
QY 642 CCATTCC 648
    |||
Db 244 CCATCCC 250
```

```
RESULT 12
US-09-907-372-5
; Sequence 5, Application US/09907372
; Patent No. US20020068242A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Warren, Bridget A.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 436
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 4048821H1
US-09-907-372-5
```

```
Query Match 25.3%; Score 164.4; DB 10; Length 436;
Best Local Similarity 99.4%; Pred. No. 1.7e-41;
Matches 165; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 CGAGTGACACTGCGAGCTACTTCTGACTGCGCGCTGCGACTGAAGCCGAGCTCAA 60
    |||
Db 136 CGAGTGACACTGCGAGCTACTTCTGACTGCGCGCTGCGACTGAAGCCGAGCTCAA 195
QY 61 AGATGAAGTTGGGAAGGTTAAACAACCACTGCTGCTCCCTGCAAGGCGAGCTTCCAGAA 120
    |||
Db 196 AGATGAAGTTGGGAAGGTTAAACAACCACTGCTGCTCCCTGCAAGGCGAGCTTCCAGAA 255
QY 121 TACCTCTCCCGCAGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCG 166
    |||
Db 256 TACCTCTCCCGCAGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCG 301
```

```
RESULT 13
US-09-907-372-14
; Sequence 14, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Warren, Bridget A.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 14
; LENGTH: 548
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 702152066H1
US-09-907-372-14
```

```
Query Match 21.7%; Score 140.6; DB 10; Length 548;
Best Local Similarity 72.5%; Pred. No. 4.4e-34;
Matches 182; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
```

```
QY 17 GAGCTACTTCTGACTGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCG 76
    |||
Db 272 GAGCGGCTTACTCTGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCG 331
QY 77 GGTAAACACCACTGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCG 136
    |||
Db 332 ACTGAAGTCAACTGTGTGCTGTAAGCCAGGACACTTCCAGAACTGCTCTCCCGCAGA 391
QY 137 GCCCGTCCAGCCCGCAGGCTGAGAGCAACCAAGTGTGAGGAGGCTCCAGGCG 196
    |||
Db 392 GCCCGTCCAGCCCGCAGGCTGAGAGCAACCAAGTGTGAGGAGGCTCCAGGCT 451
```

OY	285	GAGAGGAGACCAATCTGTACTGGAAAGCTTGGAAGCCTCCGAAGGCCATTCCATACTTCC	344
Db	365	GAGAGGGAGCCCAATCTGTACTGGAAAGCTTGGAAGCCTCCGAAGGCCATTCCATACTTCC	306
OY	345	CTGACTTGGTACACCACTGCTACCATTCTTGGAGATGTTTCCCAGTATCCACTGGGC	404
Db	305	CTGACTTGGTACACCACTGCTACCATTCTTGGAGATGTTTCCCAGTATCCACTGGGC	246
OY	405	TCCCCGACGCCCAAGTTTTGGAGCGAGGGGTGCCGCAACAGACAGTCCTCTGGACTGA	464
Db	245	TCCCCGACGCCCAAGTTTTGGAGCGAGGGGTGCCGCAACAGACAGTCCTCTGGACTGA	186
OY	465	CCAGGAGACCCGACAGTTGGAACCCGGGGGAGCAGAGCAGAGTGAGCCCATCAATGSCA	524
Db	185	CCAGGAGACCCGACAGTTGGAACCCGGGGGAGCAGAGCAGAGTGAGCCCATCAATGSCA	126
OY	525	TTCAATGTCACCGCGGGGTCTATGACTATCACTGGCAACATCTACATCTACANTGGACCG	584
Db	125	TTCAATGTCACCGCGGGGTCTATGACTATCACTGGCAACATCTACATCTACANTGGACCG	66
OY	585	TACATGGGGGAGCACACCGGGGTCTCTGGAGACCTGCCAGCTACCCCCCAACTCCATACCCA	644
Db	65	TACATGGGGGAGCACACCGGGGTCTCTGGAGACCTGCCAGCTACCCCCCAACTCCATACCCA	6
OY	645	TTCCC 	
Db	5	TTCCC 1	

RESULT 8
US-09-867-701-5601
Sequence 5601, Application US/09867701
Patent No. US20020132237A1
GENERAL INFORMATION:
APPLICANT: Agiate, Paul A.
APPLICANT: Jones, Robert
APPLICANT: Harlocker, Susan L.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.497
CURRENT APPLICATION NUMBER: US/09/867,701
CURRENT FILING DATE: 2001-05-29
NUMBER OF SEQ ID NOS: 10912
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5601
LENGTH: 281
TYPE: DNA
ORGANISM: Homo sapien
US-09-867-701-5601

	Query Match	39.1%;	Score 253.8;	DB 10;	Length 281;	
	Best Local Similarity	97.1%;	Pred. No. 2.5e-69;			
	Matches 269;	Conservative	0;	Mismatches 7;	Indels 1;	Gaps 1#;
OY	368	CCCATTTTCGGAGATGTTTTCCCCAGTATCCATCGTGGGCTCCCGGCACGCC--CAGTTTTGA	426			
Dd	5	CCACTTTTCGGAGATGTTTTCCCCAGTATCCATCGTGGGCTCCCGGCACGCCAGTTTTGA	64			
OY	427	GGCAGGGGTGCCGCAACAGCAGAGTCTCTGACCTGACCAGGGAGCCCGCACTTGGAAC	486			
Dd	65	GGCAGGGGTGCCGCAACAGCAGAGTCTCTGACCTGACCAGGGAGCCCGCACTTGGAAC	124			
OY	487	CGGGGAGCAGACCCAGGTGGCCCGAGGTACCAATGGCATTGTACCGGGGGGCTCAT	546			
Dd	125	CGGGGAGCAGACCTAGGTGGCCCGAGGTACCAATGGCATTGTACCGGGGGGCTCAT	184			
OY	547	GACTATTCATGGCAACATATACATCTACAATAAGAACAGTACGGGGGGACCAACCGGGTCC	606			
Dd	185	GACTATTCATGGCAACATATTAATCTACATATGAGCACAGTACTGGGGGGACCAACCGGGTCA	244			
OY	607	TGAGAGCTCCCGAGTACCCCGAACCCTTCATACCCC	643			

```

Db      245  TGGAGACCTCCAGCTACCCCGGAACCTCATTCCTCC 281

RESULT 9
US-09-867-701-5666/c
; Sequence 5666, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5666
; LENGTH: 289
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-5666

Query Match      38.7%; Score 251.2; DB 10; Length 289;
Best Local Similarity 98.5%; Pred. No. 1,6e-68;
Matches 264; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 379 ACATGTTTCCCGAGTATCCATCGGCTCCCGCCACCCCACTTTGGAGGAGGGGTGCC 438
      |||||||
DB 289 ACATGTTTCCCGAGTATCCATCGGCTCCCGCCACCCCACTTTGGAGGAGGGGTGCC 230
      |||||||

QY 439 GCAACAGCAGA-GTCCTCGACCTGACACGAGGAGCCGCACTTTGGAAACCCGGGAGCAGA 497
      |||||||
DB 229 GCAACAGCAGATGTCTCTCGACCTGACACGAGGAGCCGGGTGGAAACCCGGGAGCAGA 170
      |||||||

QY 498 GCCAGGTGGCCCAAGGTACCAATGGCATTCATGTACCGCGCGGCTATGACTATACACTG 557
      |||||||
DB 169 GCCAGGTGGCCCAAGGTACCAATGGCATTCATGTACCGCGCGGCTATGACTATACACTG 110
      |||||||

QY 558 GCAACATCTACATCTACATGTGACAGCAGTACTGGGGGAGACACCGGGCTCTGGAGACTCC 617
      |||||||
DB 109 GCAACATCTACATCTACATGTGACAGCAGTACTGGGGGAGACACCGGGCTCTGGAGACTCC 50
      |||||||

QY 618 CAGCTACCCCGAACCCTCCATACCCCAT 645
      |||||||
DB 49 CAGCTACCCCGAACCCTCCATTCCTCCCT 22
      |||||||

RESULT 10
US-09-867-701-4881
; Sequence 4881, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4881
; LENGTH: 481
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-4881

Query Match      32.1%; Score 208.2; DB 10; Length 481;
Best Local Similarity 93.8%; Pred. No. 4.3e-55;
Matches 228; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

```

Query Match 81.8%; Score 531; DB 10; Length 2136;
 Best Local Similarity 85.7%; Pred. No. 3e-155;
 Matches 649; Conservative 0; Mismatches 0; Indels 108; Gaps 1;

```

QY 1 CGAGTGTACACTGCGAGCTACTTTCTGTGACTGCGCCGCTGSGCACTGAAAGCGAGCTCAA 60
    |||||||
DB 579 CGAGTGTACACTGCGAGCTACTTTCTGTGACTGCGCCGCTGSGCACTGAAAGCGAGCTCAA 638
    |||||||
QY 61 AGATGAAGTGGAGGAGGTAAACAACCACTGCGTCCCTGGAAGGCGAGGCACTTCCGAA 120
    |||||||
DB 639 AGATGAAGTGGAGGAGGTAAACAACCACTGCGTCCCTGGAAGGCGAGGCACTTCCGAA 698
    |||||||
QY 121 TACCTCTCCCGCAGCGCCGCTGCCAGCCCAACAGAGTGTGAGAACCAAGGCTCGGT 180
    |||||||
DB 699 TACCTCTCCCGCAGCGCCGCTGCCAGCCCAACAGAGTGTGAGAACCAAGGCTCGGT 758
    |||||||
QY 181 GGAAGGCACTCCAGGCACTGCGCCAGTCCGACACAACTGCAAAAAATTCATTAGAGCCACT 240
    |||||||
DB 759 GGAAGGCACTCCAGGCACTGCGCCAGTCCGACACAACTGCAAAAAATTCATTAGAGCCACT 818
    |||||||
QY 241 GCGCCCAAGAGATGCA----- 256
    |||||||
DB 819 GCGCCCAAGAGATGCAAGAACCACTGATGCTGCGCTTCTGCTGCACTGCGCTTCTT 878
    |||||||
QY 257 ----- 256
    |||||||
DB 879 TCTGCTCTTGCCACCGCTTCTCTGCTGATCTGGAAGAGCAACCGCTTCTCTGCAAGAA 938
    |||||||
QY 257 -----GATTCGCTGCTCAAGAGCGCTCCGAGAGGAGAGCAACCGCTTCTGCTGCAAG 312
    |||||||
DB 939 ACTGGGATTCCTGCTCAAGAGCGCTCCGAGAGGAGAGCAACCGCTTCTGCTGCAAG 998
    |||||||
QY 313 CTGGAGCACTCCGAGAGCGCCATCCATCTTCCCTGACTGATGAGGCACTGATGAGCCAT 372
    |||||||
DB 999 CTGGAGCACTCCGAGAGCGCCATCCATCTTCCCTGACTGATGAGGCACTGATGAGCCAT 1058
    |||||||
QY 373 TTCTGAGATGTTTCCCAATATCAGTGGGCTCCCGCAGCGCCCACTTTTGGAGGCGAG 432
    |||||||
DB 1059 TTCTGAGATGTTTCCCAATATCAGTGGGCTCCCGCAGCGCCCACTTTTGGAGGCGAG 1118
    |||||||
QY 433 GGTGCCCAACAGCAGAGTCTCTGACCTGACAGGAGCGCGAGTGTGAACCGGCGGA 492
    |||||||
DB 1119 GGTGCCCAACAGCAGAGTCTCTGACCTGACAGGAGCGCGAGTGTGAACCGGCGGA 1178
    |||||||
QY 493 GCAGAGCCAGGTGGCCAGAGGTACCAATGCAATGATGATGACCGGGGCTATGACTAT 552
    |||||||
DB 1179 GCAGAGCCAGGTGGCCAGAGGTACCAATGCAATGATGATGACCGGGGCTATGACTAT 1238
    |||||||
QY 553 CACTGGCAATCTACATCTACAAATGAGCAGTACTGGGGGAGCACCGGGCTCTGGAGA 612
    |||||||
DB 1239 CACTGGCAATCTACATCTACAAATGAGCAGTACTGGGGGAGCACCGGGCTCTGGAGA 1298
    |||||||
QY 613 CCTCCAGCTACCCCGAGCACTCCATACCCCAATCCC 649
    |||||||
DB 1299 CCTCCAGCTACCCCGAGCACTCCATACCCCAATCCC 1335
    |||||||

```

RESULT 6
 US-09-907-372-8
 ; Sequence 8, Application US/09907372
 ; Patent No. US20020068242A1
 ; GENERAL INFORMATION:

```

; APPLICANT: Lal, Preeti G.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 8
; LENGTH: 574

```

TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc_feature
 OTHER INFORMATION: Incyte ID No. US20020068242A1 8234468H1
 US-09-907-372-8

Query Match 72.8%; Score 472.4; DB 10; Length 574;
 Best Local Similarity 99.8%; Pred. No. 3.6e-137;
 Matches 473; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 176 CTGGTGAAGGCACTCCAGCACTGCGCCAGTCCGACAACTGCAAAATTCATTAGAG 235
    |||||||
DB 1 CTGGTGAAGGCACTCCAGCACTGCGCCAGTCCGACAACTGCAAAATTCATTAGAG 60
    |||||||
QY 236 CCAGTGGCCCGCAGAGATGATGAGATGCTGCTCAAGGGGGTCCGCGAGAGAGGAGCC 295
    |||||||
DB 61 CCAGTGGCCCGCAGAGATGATGAGATGCTGCTCAAGGGGGTCCGCGAGAGAGGAGCC 120
    |||||||
QY 296 AATCTGTACTGGAAGCTGGAGCCTCCGAGGCCCATCACTTCCCTGACTTGTA 355
    |||||||
DB 121 AATCTGTACTGGAAGCTGGAGCCTCCGAGGCCCATCACTTCCCTGACTTGTA 180
    |||||||
QY 356 CAGCCACTGCTACCATTTCTGAGATGTTTCCCAATATCAGTGGGCTCCCGCAGCC 415
    |||||||
DB 181 CAGCCACTGCTACCATTTCTGAGATGTTTCCCAATATCAGTGGGCTCCCGCAGCC 240
    |||||||
QY 416 CCAGTTTGAAGGAGGGGTGCGCCAGACAGAGTCTGCGACCTGAGCAGGAGCGG 475
    |||||||
DB 241 CCAGTTTGAAGGAGGGGTGCGCCAGACAGAGTCTGCGACCTGAGCAGGAGCGG 300
    |||||||
QY 476 CAGTTGAACCCGGGGAGCAGAGCAGTGGGCCACGCTACCAATGCACTTCAATGTCACC 535
    |||||||
DB 301 CAGTTGAACCCGGGGAGCAGAGCAGTGGGCCACGCTACCAATGCACTTCAATGTCACC 360
    |||||||
QY 536 GGGGGTCTATGACTATCACTGCGACATCTACATCTACAAATGAGCAGTACTGGGGGA 595
    |||||||
DB 361 GGGGGTCTATGACTATCACTGCGACATCTACATCTACAAATGAGCAGTACTGGGGGA 420
    |||||||
QY 596 GCACGGGGTCTGAGACCTCCAGTACCCCGCAACCTCCATCCCAATCCC 649
    |||||||
DB 421 GCACGGGGTCTGAGACCTCCAGTACCCCGCAACCTCCATCCCAATCCC 474
    |||||||

```

RESULT 7
 US-09-907-372-9/c
 ; Sequence 9, Application US/09907372
 ; Patent No. US20020068242A1
 ; GENERAL INFORMATION:

```

; APPLICANT: Lal, Preeti G.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 425
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 7716340H1
US-09-907-372-9

```

Query Match 65.5%; Score 425; DB 10; Length 425;
 Best Local Similarity 100.0%; Pred. No. 1.8e-122;
 Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 225 ATCCATTAGAGCCACTGCCCGCAGAGATGTCAGATGCTGCTCAAAAGGCGTCCGAG 284
    |||||||
DB 425 ATCCATTAGAGCCACTGCCCGCAGAGATGTCAGATGCTGCTCAAAAGGCGTCCGAG 366
    |||||||

```

```

Db 759 GGAGGACACTCCAGGACACTGCGCCAGTCCGACACCACTGCAAAAATCATATAGGCCACT 818
QY 241 GCGCCGACAGATGCA----- 256
Db 819 GCGCCGACAGATGCAAGAACCAATGCTGAGTGGCGCTTCTGCTGCCACTGCGCTTCT 878
QY 257 ----- 256
Db 879 TCTGCTCTTGGCCACCGTCTTCTGCTGATCTGGAAGACCACTTCTCTGCGAGAA 938
QY 257 -----GATCGCTGCTCAAGAGGCGTCCGACAGGAGAGGAGCAATCCTGTAGCTGAG 312
Db 939 ACTGGGATCGCTGCTCAAGAGGCGTCCGACAGGAGAGGAGCAATCCTGTAGCTGAG 998
QY 313 CTGGGACCTCCGAGAGGCGCATTCATCTTCCCTGACTTGGTATAGGCACTGCTACCAT 372
Db 999 CTGGGACCTCCGAGAGGCGCATTCATCTTCCCTGACTTGGTATAGGCACTGCTACCAT 1058
QY 373 TTCTGAGATGTTTCCCGAGTATCCACTGGGCTCCCGCAGCCCACTTTTGGAGGAGG 432
Db 1059 TTCTGAGATGTTTCCCGAGTATCCACTGGGCTCCCGCAGCCCACTTTTGGAGGAGG 1118
QY 433 GGTGCGCCAAACAGACAGTCTCTGACCTGACCAAGGAGCGCAGTTGGAAACCCGGGGA 492
Db 1119 GGTGCGCCAAACAGACAGTCTCTGACCTGACCAAGGAGCGCAGTTGGAAACCCGGGGA 1178
QY 493 GCAGAGCAGAGTGGCCGACGATGCAATGCAATTCATGACCGGGGGGCTATGACTAT 552
Db 1179 GCAGAGCAGAGTGGCCGACGATGCAATGCAATTCATGACCGGGGGGCTATGACTAT 1238
QY 553 CACTGGCAACATCTATCATCTACATGAGACAGTACTGGGGGAGCACACCGGGTCTTGAGA 612
Db 1239 CACTGGCAACATCTATCATCTACATGAGACAGTACTGGGGGAGCACACCGGGTCTTGAGA 1298
QY 613 CTTCCAGCTACCCCGAAGCTTCATACCCCATTC 649
Db 1299 CTTCCAGCTACCCCGAAGCTTCATACCCCATTC 1335

RESULT 4
US-09-962-436-262
: Sequence 262, Application US/09962436
: Patent No. US20020081301A1
: GENERAL INFORMATION:
: APPLICANT: Soppet, Daniel
: TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
: FILE REFERENCE: 689290-75
: CURRENT APPLICATION NUMBER: US/09/962,436
: PRIOR FILING DATE: 2001-09-25
: PRIOR APPLICATION NUMBER: US/60/235,082
: PRIOR FILING DATE: 2000-09-25
: PRIOR APPLICATION NUMBER: US/60/234,924
: NUMBER OF SEQ ID NOS: 568
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 262
: LENGTH: 2136
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-962-436-262

Query Match 81.8%; Score 531; DB 10; Length 2136;
Best Local Similarity 85.7%; Pred. No. 3e-155;
Matches 649; Conservative 0; Mismatches 0; Indels 108; Gaps 1;
QY 1 CGAGGTACACACTGCGAGTACTTCTGACTGCGCGCTGCGACTGAGCGGAGCTCA 60
Db 579 CGAGGTACACACTGCGAGTACTTCTGACTGCGCGCTGCGACTGAGCGGAGCTCA 638
QY 61 AGATGAAGTTGGAGGGTAAACAACAATGCTGCTCCCTGCAAGGAGGCACTTCCAGA 120

```

```

Db 639 AGATGAAGTTGGAGGGTAAACAACAATGCTGCTCCCTGCAAGGAGGCACTTCCAGA 698
QY 121 TACCTCTCCCGACAGGCGCGCTGCGAGCCACACACAGGTGTGAGAACAAAGTCTGGT 180
Db 699 TACCTCTCCCGACAGGCGCGCTGCGAGCCACACACAGGTGTGAGAACAAAGTCTGGT 758
QY 181 GGAGGCAAGTCCAGGCACTGCGACATGTCGACACACACCTGCAAAAATCATATGAGCACT 240
Db 759 GGAGGCAAGTCCAGGCACTGCGACATGTCGACACACACCTGCAAAAATCATATGAGCACT 818
QY 241 GCGCCGACAGATGCA----- 256
Db 819 GCGCCGACAGATGCAAGAACCAATGCTGATGCTGCGCTTCTGCTGCCACTGCGCTTCT 878
QY 257 ----- 256
Db 879 TCTGCTCTTGGCCACCGTCTTCTGCTGATCTGGAAGACCACTTCTCTGCGAGAA 938
QY 257 -----GATCGCTGCTCAAGAGGCGTCCGACAGGAGAGGAGCAATCCTGTAGCTGAG 312
Db 939 ACTGGGATCGCTGCTCAAGAGGCGTCCGACAGGAGAGGAGCAATCCTGTAGCTGAG 998
QY 313 CTGGGACCTCCGAGAGGCGCATTCATCTTCCCTGACTTGGTATAGGCACTGCTACCAT 372.
Db 999 CTGGGACCTCCGAGAGGCGCATTCATCTTCCCTGACTTGGTATAGGCACTGCTACCAT 1058
QY 373 TTCTGAGATGTTTCCCGAGTATCCACTGGGCTCCCGCAGCCCACTTTTGGAGGAGG 432
Db 1059 TTCTGAGATGTTTCCCGAGTATCCACTGGGCTCCCGCAGCCCACTTTTGGAGGAGG 1118
QY 433 GGTGCGCCAAACAGACAGTCTCTGACCTGACCAAGGAGCGCAGTTGGAAACCCGGGGA 492
Db 1119 GGTGCGCCAAACAGACAGTCTCTGACCTGACCAAGGAGCGCAGTTGGAAACCCGGGGA 1178
QY 493 GCAGAGCAGAGTGGCCGACGATGCAATGCAATTCATGACCGGGGGGCTATGACTAT 552
Db 1179 GCAGAGCAGAGTGGCCGACGATGCAATGCAATTCATGACCGGGGGGCTATGACTAT 1238
QY 553 CACTGGCAACATCTATCATCTACATGAGACAGTACTGGGGGAGCACACCGGGTCTTGAGA 612
Db 1239 CACTGGCAACATCTATCATCTACATGAGACAGTACTGGGGGAGCACACCGGGTCTTGAGA 1298
QY 613 CTTCCAGCTACCCCGAAGCTTCATACCCCATTC 649
Db 1299 CTTCCAGCTACCCCGAAGCTTCATACCCCATTC 1335

RESULT 5
US-09-880-107-2135
: Sequence 2135, Application US/09880107
: Patent No. US20020142981A1
: GENERAL INFORMATION:
: APPLICANT: Horne, Darci T.
: APPLICANT: Vockley, Joseph G.
: APPLICANT: Scherf, Uwe
: TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
: FILE REFERENCE: 44921-5028-WO
: CURRENT APPLICATION NUMBER: US/09/880,107
: PRIOR FILING DATE: 2001-06-14
: PRIOR APPLICATION NUMBER: US 60/211,379
: PRIOR FILING DATE: 2000-06-14
: PRIOR APPLICATION NUMBER: US 60/237,054
: NUMBER OF SEQ ID NOS: 3950
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 2135
: LENGTH: 2136
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: Genbank Accession No. US20020142981A1 I04270
US-09-880-107-2135

```

Query	March	Similarity	92.6%	Score	600.8	DB	10	Length	651
Best Local	Similarity	99.2%	Pred.	No. 4.7e-177					
Matches	635	Conservative	0	Mismatches	2	Indels	3	Gaps	3

QY	13	CTGCAGGCTACTTTCCTGACTGCCCGCCCTGCGCACTTAAGCCGAGCTCAAAAGATGAAGTTGG	72
Db	640	CTGGAGGCTACTTTCCTGACTGCCCGCCCTGCGCACTTAAGCCGAGCTCAAAAGATGAAGTTGG	581
QY	73	GAAGGGTAAACAACCACTGGCTCCCTGCGAAGGCGAGGGCACTTCCAGATACCTCTCTCCC	132
Db	580	GAAGGGTAAACAACCACTGGCTCCCTGCGAAGGCGAGGGCACTTCCAGATACCTCTCTCCC	521
QY	133	CAGGCCCC-GCGGCGAGGCCCAACAGAGGTGGAACCAACAGGTGTGGTGGAGGGAGGTC	191
Db	520	CAGGCCCCCTGGCTCGCAGGCCCAACAGAGGTGGAACCAACAGGTGTGGTGGAGGGAGGTC	461
QY	192	CAGGCACTGCCAGTCCGACACAACTGTGAAAAATTCATTAGAGCACTGCCCCAGAGA	251
Db	460	CAGGCACTGCCAGTCCGACACAACTGTGAAAAATTCATTAGAGCACTGCCCCAGAGA	401
QY	252	TGTCAAGATCGTGTCAAGAGGCGTCCGACAGGAGAGGAGCAATCCTGTAGCTGAA	311

Query Match	Similarity	81.8%	Score 531	DB 9	Length 2136
Best Local	Similarity	85.7%	Pred. No. 3e-155		
Matches	649	Conservative	0	Mismatches	0
				Indels	108
				Gaps	1
QY	1	CGAGTGCACACACTGCGAGCTACTTCTTGACTGCGCCGCTGCGACTGAAGCCGAGCTCAA	60		
Db	579	CGAGTGCACACACTGCGAGCTACTTCTTGACTGCGCCGCTGCGACTGAAGCCGAGCTCAA	638		
QY	61	AGATGAAGTTGGGAAGGGTAACACACCCTGCTCCCCCTGCACAGGCGACATTTCACAGA	120		
Db	639	AGATGAAGTTGGGAAGGGTAACACACCCTGCTCCCCCTGCACAGGCGACATTTCACAGA	698		
QY	121	TACCTCCTCCGCCACGCGCCGCTGCGACGCCCCACACACAGGTGTGAAGAACCAAGTCTGGT	180		
Db	699	TACCTCCTCCGCCACGCGCCGCTGCGACGCCCCACACACAGGTGTGAAGAACCAAGTCTGGT	758		
QY	181	GGAGCGAGCTCCAGGCGACTGCTCCGAGTCCGACCAACACTGCAAAAAATCAATTGAGCCACT	240		

OY	1	CAGGTGATACACACACACGCCGGCTACTTTCTTGACTTCCCGCCTTCACCTTAAGGCACCACTGAA	
Dd	629	CGAGGTACACACACTGCCAGCTACTTTCTGACTTGCCTCCCTGGCATGAAGCCGAGCTCAA	6888
OY	61	AGATGAAGTTGGGAAGGTAACAACCACACTGGTGTCCCTCGCAAGGCAGAAGGCACTTCCAGAA	120
Dd	689	AGATGAAGTTGGGAAGGTAACAACCACACTGGTGTCCCTCGCAAGGCAGAAGGCACTTCCAGAA	748
OY	121	TACCTCCCTCCCCCAGCGCGCGCTGCGACGCCCATCACCAAGGTGTGAGAACCAAGGTCTGGT	180
Dd	749	TACCTCCCTCCCCCAGCGCGCGCTGCGACGCCCATCACCAAGGTGTGAGAACCAAGGTCTGGT	808
OY	141	GSAGGCACCTCCAGGCACATTGGCCACAGTCCGACACAACCTGCATAAATCATATTAGAGCCACT	240
Dd	809	GGAGGCACCTCCAGGCACATTGGCCACAGTCCGACACAACCTGCATAAATCATATTAGAGCCACT	868


```
QY      80 AACACCACTGCGTCCCTGCAAGGAGGCGACTTCAGAAATACCTCCTCCCGGCGCC 139
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      756 AACACGAGATGCCCGGTGGCCCGCCAGGACACCTTCTCAGCAGCAGCTCCAGCTCAGAG 815
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      140 CGCTGCCAGCCCCACACCACTGTGAGAACCAAGCTCTGTGGAGGACGCTCCAGGCACT 199
      | ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      816 CAGTGCAGGCCCCACGCACTGCACGCGCCTGGGCTGGCCCTCAATGTGCGCAGGCTCT 875
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      200 GCCCAGTCCGACACACCTGCA 221
      || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      876 TCCTCCCATGACACCTGTGCA 897
      || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 15

```
US-09-286-529-19
; Sequence 19, Application US/09286529
; Patent No. 6297367
; GENERAL INFORMATION:
; APPLICANT: Catharine Tribouley
; TITLE OF INVENTION: NEW MEMBERS OF TNF AND TNFR FAMILIES
; FILE REFERENCE: 1408.003/200130.439C1
; CURRENT APPLICATION NUMBER: US/09/286,529
; CURRENT FILING DATE: 1999-04-05
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 1859
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-286-529-19
```

Query Match

5.6%; Score 36.4; DB 4; Length 1859;

Best Local Similarity 53.5%; Pred. No. 0.54; Mismatches 66; Indels 0; Gaps 0;

```
Matches 76; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY      80 AACACCACTGCGTCCCTGCAAGGAGGCGACTTCAGAAATACCTCCTCCCGGCGCC 139
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      516 AACACGAGATGCCCGGTGGCCCGCCAGGACACCTTCTCAGCAGCAGCTCCAGCTCAGAG 575
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      140 CGCTGCCAGCCCCACACCACTGTGAGAACCAAGCTCTGTGGAGGACGCTCCAGGCACT 199
      | ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      576 CAGTGCAGGCCCCACGCACTGCACGCGCCTGGGCTGGCCCTCAATGTGCGCAGGCTCT 635
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      200 GCCCAGTCCGACACACCTGCA 221
      || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      636 TCCTCCCATGACACCTGTGCA 657
      || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: April 15, 2003, 16:28:35
Job time : 29.9124 secs

SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1004
TYPE: DNA
ORGANISM: Homo sapiens
US-08-114-944D-1

Query Match
Best Local Similarity 56.1%; Pred. No. 0.074;
Matches 74; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Y 79 TAACACCACTGCGTCCCTGCAAGCAGGCGCACTTCAGAAATACCTCCGCCAGGCG 138
D 464 TGATACACATCTGCGAGCGCTCCAGTCGGCTTCTTCCAAATGTGATCTGCTTTCGA 523
Y 139 CCGCTGCAGCCGCCACACACAGAGTGTGAGAACCAAGTGTGTGGAGCGAGCTCCAGGCAC 198
D 524 AAAATGTCACCTTGGAAGAGCTGTGAGACCAAGACCTGTGTGCAACAGCAGGCGAC 583
Y 199 TGCCCAAGTCGA 210
D 584 AACACAGCTCA 595

RESULT 12
US-09-286-529-7
Sequence 7, Application US/09286529
Patent No. 6297367
GENERAL INFORMATION:
APPLICANT: Catheline Tribouley
TITLE OF INVENTION: NEW MEMBERS OF TNF AND TNFR FAMILIES
FILE REFERENCE: 1408.003/200130.439C1
CURRENT FILING DATE: US/09/286,529
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7
LENGTH: 459
TYPE: DNA
ORGANISM: human
US-09-286-529-7

Query Match
Best Local Similarity 53.5%; Pred. No. 0.35;
Matches 76; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

Y 80 AACACCACTGCGTCCCTGCAAGCAGGCGCACTTCAGAAATACCTCCGCCAGGCGC 139
D 187 AACACGAGTGCAGCGCTCCGCCAGGCGCACTTCAGAAATACCTCCGCCAGGCGC 246
Y 140 CGCTGCAGCCGCCACACACAGAGTGTGAGAACCAAGTGTGTGGAGCGAGCTCCAGGCAC 199
D 247 CAGTGCAGCGCCGCCACACAGAGTGTGAGAACCAAGTGTGTGGAGCGAGCTCCAGGCAC 306
Y 200 GCCCAAGTCCGACACAACTGCA 221
D 307 TCCTCCCATGACACCTGTGCA 328

RESULT 13
US-08-794-796-1
Sequence 1, Application US/08794796
Patent No. 5885800
GENERAL INFORMATION:
APPLICANT: Emery, John
APPLICANT: Tan, KB
APPLICANT: Truneh, Alem
APPLICANT: Young, Peter
TITLE OF INVENTION: Tumor Necrosis Related Receptor,
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation

STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/794,796
FILING DATE: 04-FEB-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Han, William T
REGISTRATION NUMBER: 34,344
REFERENCE/DOCKET NUMBER: GH50000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5219
TELEFAX: 610-270-4026
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1164 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-794-796-1

Query Match
Best Local Similarity 53.5%; Pred. No. 0.47;
Matches 76; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

Y 80 AACACCACTGCGTCCCTGCAAGCAGGCGCACTTCAGAAATACCTCCGCCAGGCGC 139
D 549 AACACGAGTGCAGCGCTCCGCCAGGCGCACTTCAGAAATACCTCCGCCAGGCGC 608
Y 140 CGCTGCAGCCGCCACACACAGAGTGTGAGAACCAAGTGTGTGGAGCGAGCTCCAGGCAC 199
D 609 CAGTGCAGCGCCGCCACACAGAGTGTGAGAACCAAGTGTGTGGAGCGAGCTCCAGGCAC 668
Y 200 GCCCAAGTCCGACACAACTGCA 221
D 669 TCCTCCCATGACACCTGTGCA 690

RESULT 14
US-09-286-529-18
Sequence 18, Application US/09286529
Patent No. 6297367
GENERAL INFORMATION:
APPLICANT: Catheline Tribouley
TITLE OF INVENTION: NEW MEMBERS OF TNF AND TNFR FAMILIES
FILE REFERENCE: 1408.003/200130.439C1
CURRENT FILING DATE: US/09/286,529
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 18
LENGTH: 1347
TYPE: DNA
ORGANISM: Homo sapien
US-09-286-529-18

Query Match
Best Local Similarity 53.5%; Pred. No. 0.49;
Matches 76; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

```

?      LENGTH: 519 base pairs
?      TYPE: nucleic acid
?      STRANDEDNESS: single
?      TOPOLOGY: linear
?      MOLECULE TYPE: cDNA
?      HYPOTHETICAL: NO
?      ANTI-SENSE: NO
?      ORIGINAL SOURCE:
?      ORGANISM: HUMAN
?      IMMEDIATE SOURCE:
?      CLONE: CD40 EXTRACELLULAR REGION
?      US-09-399-106-4

```

Query Match	6.08:	Score 39.2:	DB 4:	Length 519:
Best Local Similarity	56.18:	Pred. No. 0.06:		
Matches 74:	Conservative 0:	Mismatches 58:	Indels 0:	Gaps 0

QY	79	TAACAAOCACGCGCGTCCCGTCGACAGGCGACGCACTTCCAGAAATACCTTCCTCCCGCAGCG	138
Db	359	TGAATACCACTCTGCGAGCGCCGCGCCAGTCGGGTCTTTCCAAATGTGTCAATCTGTGTTTCGA	418
QY	139	CCGCGCCACGCCACACACACAGCTGTGAGAACCAAGCTCTGTGAGGCGACACTCCAGGCAC	198
Db	419	AAAAATGTCACCTTGTGGACAAACGTCGTGAGACCAAGACCTGTTGTGTGCACAGGAGGACAC	478
QY	199	TGCCCAAGTCGA	210
Db	479	AAACCAAGACTGA	490

RESULT 9
US-09-071-433-85
; Sequence 85, Application. US/09071433A

```

1  APPLICANT: Bennett, C. Frank
2  APPLICANT: Cowser, Lex M
3  TITLE OF INVENTION: Antisense Modulation of CD40 Expression
4  FILE REFERENCE: RTS-0002
5  CURRENT APPLICATION NUMBER: US/09/071,433A
6  CURRENT FILING DATE: 1998-05-01
7  NUMBER OF SEQ ID NOS: 91
8  SOFTWARE: PatentIn Ver. 2.0
9  SEQ ID NO 85
10 LENGTH: 1004
11 TYPE: DNA
12 ORGANISM: Artificial Sequence
13 FEATURE:
14 OTHER INFORMATION: Description of Artificial Sequence:Synthetic
15 US-09-071-433-B5

```

Query Match	6.08	Score 39.2	DB 4	Length 1004
Best local Similarity	56.18	Pred. No. 0.074		
Matches 74	Conservative 0	Mismatches 58	Indels 0	Gaps 0

QY	79	TAAACACCATCGCTCCCTCCGCAAGGAGGGCACTTCCAGAAATACCTCCCTCCCAAGGC	138
Db	464	TGATATCCATCTGCGAGAGCCCTCCAGTCCGCTTCTTCCATATGTCAATCTGCTTTGCA	523
QY	139	CCGCTGCGACCCCAACACACAGGTGTGAGAAACCAAGTCTGTGTGAGGACACTCCAGGCAC	198
Db	524	AAAAATGTCACCTCTGGACACAGAGCTGTGAGACCAAAAGACCTGTGTGTGCAACAGGACGACAC	583
QY	199	TGCCCAAGTCGA	210
Db	584	AAACAGAGACTGA	595

RESULT 10
US-09-041-886-26
; Sequence 26, Application US/09041886
; Patent No. 6235872
; GENERAL INFORMATION:

```

1  APPLICANT:  Bresden, Dale E.
2  APPLICANT:  Rabizadeh, Sharroz
3  TITLE OF INVENTION:  Propeptotic Peptides, Dependence
4  TITLE OF INVENTION:  Polypeptides and Methods of Use
5  NUMBER OF SEQUENCES:  72
6  CORRESPONDENCE ADDRESS:
7  ADDRESSEE:  Campbell & Flores LLP
8  STREET:  4370 La Jolla Village Drive, Suite 700
9  CITY:  San Diego
10 STATE:  California
11 COUNTRY:  United States
12 ZIP:  92122
13
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE:  Floppy disk
16 COMPUTER:  IBM PC compatible
17 OPERATING SYSTEM:  PC-DOS/MS-DOS
18 SOFTWARE:  PatentIn Release #1.0, Version #1.25
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER:  US/09/041,886
21 FILING DATE:
22
23 CLASSIFICATION:
24 ATTORNEY/AGENT INFORMATION:
25 NAME:  Campbell, Cathryn A.
26 REGISTRATION NUMBER:  31,815
27 REFERENCE/DOCKET NUMBER:  P-LJ 2626
28 TELECOMMUNICATION INFORMATION:
29 TELEPHONE:  (619) 535-9001
30 TELEFAX:  (619) 535-8949
31 INFORMATION FOR SEQ ID NO:  26:
32 SEQUENCE CHARACTERISTICS:
33 LENGTH:  1004 base pairs
34 TYPE:  nucleic acid
35 STRANDEDNESS:  single
36 TOPOLOGY:  linear
37 MOLECULE TYPE:  DNA (genomic)
38 FEATURE:
39 NAME/KEY:  CDS
40 LOCATION:  48..876
41
42 OS-09-041-886-26

```

[illegible]

RESULT 11
 US-08-114-944D-1
 : Sequence 1, Application US/08114944D
 : Patent No. 6376459
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Aruffo, Alejandro J
 : APPLICANT: Ledbetter, Jeffrey A
 : APPLICANT: Stamenkovic, Ivan
 :
 : APPLICANT: No. 63764591le, Randolph
 :
 : TITLE OF INVENTION: THE CD40CR RECEPTOR AND LIGANDS THEREFOR
 :
 : FILE REFERENCE: 5624-232-999
 :
 : CURRENT APPLICATION NUMBER: US/08/114,944D
 :
 : CURRENT FILING DATE: 1993-08-31
 :
 : PRIOR APPLICATION NUMBER: 07/835,799
 :
 : PRIOR FILING DATE: 1992-02-14
 :
 : NUMBER OF SEQ ID NOS: 3

QY	199	TGCCAGTCCGA	210
Db	479	AAACAAGACTGA	490

```

RESULT 7
US-08-770-981-4
: Sequence 4, Application US/08770981
: Patent No. 6391657
: GENERAL INFORMATION:
: APPLICANT: ARMITAGE, RICHARD
: APPLICANT: FANSLAW, WILLIAM
: APPLICANT: SPRINGS, MELANIE
: APPLICANT: SRINIVASAN, SUBHASHINI
: APPLICANT: GIBSON, MARLOU
: TITLE OF INVENTION: NOVEL CYTOKINE THAT BINDS CD40
: NUMBER OF SEQUENCES: 24
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: IMMUNEX CORPORATION
: STREET: 51 UNIVERSITY STREET
: CITY: SEATTLE
: STATE: WASHINGTON
: COUNTRY: USA
: ZIP: 98101
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: Apple Macintosh
: OPERATING SYSTEM: Apple Operating System 7.1
: SOFTWARE: Microsoft Word for Apple, version 5.1a
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/770,981
: FILING DATE: 20-DEC-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/477,733
: FILING DATE: 07-JUN-1995
: APPLICATION NUMBER: 08/249,189
: FILING DATE: May 24, 1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/969,703
: FILING DATE: October 23, 1992
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/805,723
: FILING DATE: December 5, 1991
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/783,707
: FILING DATE: October 25, 1991
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Perkins, Patricia A.
: REGISTRATION NUMBER: 34,693
: REFERENCE/DOCKET NUMBER: 2802-D
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 2065870430
: TELEFAX: 2065870606
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 519 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHEetical: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: HUMAN
: IMMEDIATE SOURCE:
: CLONE: CD40 EXTRACELLULAR REGION
: IS-08-770-981-4

```

Query Match	6.0%	Score 39.2	DB 4	Length 519
Best Local Similarity	56.1%	Pred. No. 0.06		
Matches 74	Conservative 0	Mismatches 58	Indels 0	Gaps 0
QY	79	TAAACACCACTGCGTCCCTGCAGGAGGAGGACTCTCCAGAAATACCTCCTCCCGACGCG	138	
Db	359	TGATACAMCTGCGAGCCCTGCCAGTGGCGCTTCTTCACATGTGTATCTGGCTTTGCA	418	
QY	139	CCGCTGCCAGGCCCCACACCAGGTGTGAGAACCAAGTCTGGTGGAGCGAGCTCCAGGCAC	198	
Db	419	AAAATGTACACCTTGAGCAAGCTGTGAGACCAAGAACCTGGTTGTGCAACAGCAGGCGAC	478	
QY	199	TGCCCACTCGCA	210	
Db	479	AAACAGACTGA	490	

RESULT 8
 US-09-399-106-4
 Sequence 4, Application US/09399106
 Patent No. 6410711
 GENERAL INFORMATION:
 APPLICANT: ARMITAGE, RICHARD
 APPLICANT: FANSLAW, WILLIAM
 APPLICANT: SPRIGGS, MELANIE
 APPLICANT: SRINIVASAN, SUBHASHINI
 APPLICANT: GIBSON, NARAYLOU
 APPLICANT: MORRIS, ARVIA E.
 APPLICANT: MCGREW, JEFFERY
 TITLE OF INVENTION: NOVEL CYTOKINE THAT BINDS CD40
 NUMBER OF SEQUENCES: 26
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: IMMUNE CORPORATION
 STREET: 51 UNIVERSITY STREET
 CITY: SEATTLE
 STATE: WASHINGTON
 COUNTRY: USA
 ZIP: 98101
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: Apple Macintosh
 OPERATING SYSTEM: Apple Operating System 7.5.5
 SOFTWARE: MS word for Apple Power Macintosh, version 6.0.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/399,106
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/477,733
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/969,703
 FILING DATE: October 23, 1992
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/805,723
 FILING DATE: December 5, 1991
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/783,707
 FILING DATE: October 25, 1991
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Perkins, Patricia A.
 REGISTRATION NUMBER: 34,693
 REFERENCE/DOCKET NUMBER: 2802-D
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 2065870430
 TELEFAX: 2065870606
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:

1 TITLE OF INVENTION: NOVEL CYTOKINE THAT BINDS CD40
2 NUMBER OF SEQUENCES: 24
3 CORRESPONDENCE ADDRESS:
4 ADDRESSEE: IMMUNEX CORPORATION
5 STREET: 51 UNIVERSITY STREET
6 CITY: SEATTLE
7 STATE: WASHINGTON
8 COUNTRY: USA
9 ZIP: 98101
10
11 COMPUTER READABLE FORM:
12 MEDIUM TYPE: Floppy disk
13 COMPUTER: Apple Macintosh
14 OPERATING SYSTEM: Apple Operating System 7.1
15 SOFTWARE: Microsoft Word for Apple, version 5.1a
16 CURRENT APPLICATION DATA:
17 APPLICATION NUMBER: US/08/769,819
18 FILING DATE: 19-DEC-1996
19 CLASSIFICATION: 424
20
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER: 08/484,624
23 FILING DATE: 07-JUN-1995
24 APPLICATION NUMBER: 08/249,189
25 FILING DATE: May 24, 1994
26
27 PRIOR APPLICATION DATA:
28 APPLICATION NUMBER: 07/969,703
29 FILING DATE: October 23, 1992
30 APPLICATION NUMBER: 07/805,723
31 FILING DATE: December 5, 1991
32
33 PRIOR APPLICATION DATA:
34 APPLICATION NUMBER: 07/783,707
35 FILING DATE: October 25, 1991
36 ATTORNEY/AGENT INFORMATION:
37 NAME: Perkins, Patricia A.
38 REGISTRATION NUMBER: 34,693
39 REFERENCE/DOCKET NUMBER: 2802-E
40 TELECOMMUNICATION INFORMATION:
41 TELEPHONE: 2065870430
42 TELEFAX: 2065870606
43
44 INFORMATION FOR SEQ ID NO: 4:
45 SEQUENCE CHARACTERISTICS:
46 LENGTH: 519 base pairs
47 TYPE: nucleic acid
48 STRANDEDNESS: single
49 TOPOLOGY: linear
50 MOLECULE TYPE: cDNA
51 HYPOTHETICAL: NO
52 ANTI-SENSE: NO
53 ORIGINAL SOURCE:
54 ORGANISM: HUMAN
55 IMMEDIATE SOURCE:
56 CLONE: CD40 EXTRACELLULAR REGION
57
58 US-08-769-819-4

Query Match 6.0%; Score 39.2; DB 4; Length 519;
Best Local Similarity 56.1%; Pred. No. 0.06;
Matches 74; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
Db 79 TAACACCACTGCTCCCTGCAAGGAGGCACTTCAGATACCTCTCCCGCAGCGC 138
139 CCGCTGCAGCCCAACACAGGTGTGAGAACCAAGTGTGAGGAGAGTCCAGGAC 198
419 AAATGTCACCTTGCAAGAGCTGTGAGACCAAGACCTGTTGTGCAACAGGAGCGAC 478
QY 199 TGCCAGTCCGA 210
Db 479 AAACAAGACTGA 490

RESULT 6
US-08-770-974-4

1 Sequence 4, Application US/08770974
2 Patent No. 6290972
3 GENERAL INFORMATION:
4 APPLICANT: ARMITAGE, RICHARD
5 APPLICANT: FANSLON, WILLIAM
6 APPLICANT: SPRIGGS, MELANIE
7 APPLICANT: SRINIVASAN, SUBHASHINI
8 APPLICANT: GIBSON, MARYLOU
9 TITLE OF INVENTION: NOVEL CYTOKINE THAT BINDS CD40
10 NUMBER OF SEQUENCES: 24
11 CORRESPONDENCE ADDRESS:
12 ADDRESSEE: IMMUNEX CORPORATION
13 STREET: 51 UNIVERSITY STREET
14 CITY: SEATTLE
15 STATE: WASHINGTON
16 COUNTRY: USA
17 ZIP: 98101
18
19 COMPUTER READABLE FORM:
20 MEDIUM TYPE: Floppy disk
21 COMPUTER: Apple Macintosh
22 OPERATING SYSTEM: Apple Operating System 7.1
23 SOFTWARE: Microsoft Word for Apple, version 5.1a
24 CURRENT APPLICATION DATA:
25 APPLICATION NUMBER: US/08/770,974
26 FILING DATE: 20-DEC-1996
27 CLASSIFICATION: 435
28
29 PRIOR APPLICATION DATA:
30 APPLICATION NUMBER: 08/477,733
31 FILING DATE: 02-AUG-1995
32 APPLICATION NUMBER: 08/249,189
33 FILING DATE: May 24, 1994
34
35 PRIOR APPLICATION DATA:
36 APPLICATION NUMBER: 07/969,703
37 FILING DATE: October 23, 1992
38 APPLICATION NUMBER: 07/805,723
39 FILING DATE: December 5, 1991
40
41 PRIOR APPLICATION DATA:
42 APPLICATION NUMBER: 07/783,707
43 FILING DATE: October 25, 1991
44 ATTORNEY/AGENT INFORMATION:
45 NAME: Perkins, Patricia A.
46 REGISTRATION NUMBER: 34,693
47 REFERENCE/DOCKET NUMBER: 2802-D
48 TELECOMMUNICATION INFORMATION:
49 TELEPHONE: 2065870430
50 TELEFAX: 2065870606
51
52 INFORMATION FOR SEQ ID NO: 4:
53 SEQUENCE CHARACTERISTICS:
54 LENGTH: 519 base pairs
55 TYPE: nucleic acid
56 STRANDEDNESS: single
57 TOPOLOGY: linear
58 MOLECULE TYPE: cDNA
59 HYPOTHETICAL: NO
60 ANTI-SENSE: NO
61 ORIGINAL SOURCE:
62 ORGANISM: HUMAN
63 IMMEDIATE SOURCE:
64 CLONE: CD40 EXTRACELLULAR REGION
65
66 US-08-770-974-4

Query Match 6.0%; Score 39.2; DB 4; Length 519;
Best Local Similarity 56.1%; Pred. No. 0.06;
Matches 74; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
Db 79 TAACACCACTGCTCCCTGCAAGGAGGCACTTCAGATACCTCTCCCGCAGCGC 138
139 CCGCTGCAGCCCAACACAGGTGTGAGAACCAAGTGTGAGGAGAGTCCAGGAC 198
419 AAATGTCACCTTGCAAGAGCTGTGAGACCAAGACCTGTTGTGCAACAGGAGCGAC 478
QY 199 TGCCAGTCCGA 210
Db 479 AAACAAGACTGA 490

FILED DATE: May 24, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/969,703
FILING DATE: October 23, 1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/805,723
FILING DATE: December 5, 1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/783,707
FILING DATE: October 25, 1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia A.
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2802-D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 2065870430
TELEFAX: 2065870606
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 519 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE: HUMAN
ORGANISM: HUMAN
IMMEDIATE SOURCE:
CLONE: CD40 EXTRACELLULAR REGION
US-08-477-733B-4

Query Match
Best Local Similarity 56.1%; Score 39.2; DB 2; Length 519;
Pred. No. 0.06;
Matches 74; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 79 TACACCACTGCTGCTCCCTGCAAGCAGGCGCTTCAGATCTCTCCCGCCAGCGC 138
DB 359 TGATACCACTGCGAGCGAGCGCTTCTTCCAAATGCTCATCTGCTTTCGA 418
QY 139 CCGCTCCAGCCGCCACACAGGTGTGAGAACCAAGTGTGTGAGAGCGAGCTCCAGCGC 198
DB 419 AAATGTCAACCTGTGAGCAAGCTGTGAGACCAAGACCTGTGTGCAACAGCGAGCGC 478
QY 199 TGCCCGAGTCCGA 210
DB 479 AACCAAGACTGA 490

RESULT 4
US-09-088-913A-4
Sequence 4, Application US/09088913A
Patent No. 6087329
GENERAL INFORMATION:
APPLICANT: ARMITAGE, RICHARD
APPLICANT: FANSLAW, WILLIAM
APPLICANT: SPRIGGS, MELANIE
APPLICANT: SRINIVASAN, SUBHASHINI
APPLICANT: GIBSON, MARYLOU
APPLICANT: MORRIS, ARVIA E.
APPLICANT: MCGREW, JEFFERY
TITLE OF INVENTION: NOVEL CYTOKINE THAT BINDS CD40
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMUNEX CORPORATION
STREET: 51 UNIVERSITY STREET
CITY: SEATTLE
STATE: WASHINGTON
COUNTRY: USA

ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: MS Word for Apple Power Macintosh, version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/088,913A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/484,624
FILING DATE:
APPLICATION NUMBER: 08/477,733
FILING DATE: June 07, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/969,703
FILING DATE: October 23, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/805,723
FILING DATE: December 5, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/783,707
FILING DATE: October 25, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia A.
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2802-D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 2065870430
TELEFAX: 2065870606
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 519 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE: HUMAN
ORGANISM: HUMAN
IMMEDIATE SOURCE:
CLONE: CD40 EXTRACELLULAR REGION
US-09-088-913A-4

Query Match
Best Local Similarity 56.1%; Score 39.2; DB 3; Length 519;
Pred. No. 0.06;
Matches 74; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 79 TACACCACTGCTGCTCCCTGCAAGCAGGCGCTTCAGATCTCTCCCGCCAGCGC 138
DB 359 TGATACCACTGCGAGCGAGCGCTTCTTCCAAATGCTCATCTGCTTTCGA 418
QY 139 CCGCTCCAGCCGCCACACAGGTGTGAGAACCAAGTGTGTGAGAGCGAGCTCCAGCGC 198
DB 419 AAATGTCAACCTGTGAGCAAGCTGTGAGACCAAGACCTGTGTGCAACAGCGAGCGC 478
QY 199 TGCCCGAGTCCGA 210
DB 479 AACCAAGACTGA 490

RESULT 5
US-08-769-819-4
Sequence 4, Application US/08769819
Patent No. 6264951
GENERAL INFORMATION:
APPLICANT: ARMITAGE, RICHARD
APPLICANT: FANSLAW, WILLIAM
APPLICANT: SPRIGGS, MELANIE
APPLICANT: SRINIVASAN, SUBHASHINI
APPLICANT: GIBSON, MARYLOU

```

?      TYPE: nucleic acid
?      STRANDEDNESS: single
?      TOPOLOGY: linear
?      MOLECULE TYPE: cDNA
?      HYPOTHETICAL: NO
?      AMPLI-SENSE: NO
?      ORIGINAL SOURCE:
?      ORGANISM: HUMAN
?      IMMEDIATE SOURCE:
?      CLONE: CD40 EXTRACELLULAR REGION
?      US-08-249-189-4

```

Query Match	6.0%;	Score 39.2;	DB 2;	Length 519;
Best Local Similarity	56.1%;	Pred. No. 0.06;		
Matches 74;	Conservative 0;	Mismatches 58;	Indels 0;	Gaps 0;

[illegible]

QY	199	TGCCAGTCCGA	210
Db	479	AAACAAGACTGA	490

RESULT 2
US-08-484-624A-4
: Sequence 4, Application US/08484624A

GENERAL INFORMATION:

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

ADDRESSEE: IMMUNEX CORPORATION
STREET: 51 UNIVERSITY STREET
CITY: SEATTLE
STATE: WASHINGTON
COUNTRY: USA
ZIP: 98101

```

1  :
2  : COMPUTER READABLE FORM:
3  : MEDIUM TYPE: Floppy disk
4  :
5  : COMPUTER: Apple Macintosh
6  :
7  : OPERATING SYSTEM: Apple Operating System 7.5.5
8  :
9  : SOFTWARE: MS Word for Apple Power Macintosh, version 6.0.1
10 :
11 : CURRENT APPLICATION DATA:
12 :
13 : APPLICATION NUMBER: US/08/484,624A
14 :

```

CLASSIFICATION:
PRIOR APPLICATION:

;; APPLICATION NUMBER: 08/477,733
;; FILING DATE: June 07, 1995
;;

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/060,703

APPLICATIION NUMBER: 01/909,103
FILING DATE: October 23, 1992
CLASSIFICATION:

;
 ;
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/805,723
 ;

FILING DATE: December 5, 1991
CLASSIFICATION:

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/783,707

```

1 FILING DATE: October 25, 1991
2
3 CLASSIFICATION:
4
5 ATTORNEY/AGENT INFORMATION:
6
7 NAME: Perkins, Patricia A.
8
9 REGISTRATION NUMBER: 34,693
10
11 REFERENCE/DOCKET NUMBER: 2802
12
13 TELECOMMUNICATION INFORMATION:
14
15 TELEPHONE: 20658760430
16
17 TELEFAX: 2065870606
18
19 INFORMATION FOR SDO, ID NO.: 4:
20
21 SOURCE CHARACTERISTICS:

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:

SEQUENCE CHARACTERISTICS:
LENGTH: 519 base pairs

TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
 MOLECULE TYPE: CDNA

HYPOTHETICAL: NO
 PAYMENT-SENSE: NO

ANTI-SENSE: NO
ORIGINAL SOURCE:

ORGANISM: HUMAN
IMMEDIATE SOURCE:

CLONE: CD40 EXTRACELLULAR
-08-484-624A-4

Answer Match

Query Match	6.08;	Score	39.2;	DB	2;	Length	519;
Best Local Similarity	56.18;	Pred. No.	0.06;				
Matches	74;	Conservative	0;	Mismatches	58;	Indels	0;
						Gaps	0;

Qy 79 TACCAACCACCTGGCGTCCCTGCAGGCGAGGGACTTCAGAATACCTCCTCCCCCAGAGGC 138
| | | | | | | | | | | | | | | |
Db 359 TGATTACCATCTGCAGGCCCTGGCCAGTCGGCTTCTTCCAATGTGTCAATCTGCTTTGA 418

Qy 139 CCGCTGCCACCCCCCACCAGGTGTGAGAACCAGGTCGTGGAGGCAGCTCCAGGCAC 198
|| ||| | | | | | | | | | | | | |
Db 419 AAAATGTCAACCTTGGACAACACTGTGAGACCAAAGACCTGGTTTGCAACAGGCAGGCAC 478

Qy	199	TGCCAGTCCGA	210
Db	479	AAACAAGACTGA	490

RESULT 3
US-08-477-733B-4
; Sequence 4, Application US/08477733B

GENERAL INFORMATION:

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

CITY: SEATTLE
STATE: WASHINGTON

COUNTRY: USA
ZIP: 98101

```

; COMPUTER READABLE FORM
; MEDIUM TYPE: FLOPPY
; GOVERNMENT ACQUISITION

```

```

;
; COMPUTER: Apple M
; OPERATING SYSTEM:
; SOFTWARE: MS Word

```

```

; COLIMAKE. NO MORE
; CURRENT APPLICATION
; APPLICATION NUMBER

```

; FILING DATE: June
 ; CLASSIFICATION: 4
 ;

; PRIOR APPLICATION DATA
; APPLICATION NUMBER
;

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:11:20 ; Search time 20.9124 Seconds
(without alignments)
9517.464 Million cell updates/sec

Title: US-09-917-372-2_COPY_629_1277

Perfect score: 649
Sequence: 1 cgagtgtaacactgcgagc.....aacctcaccattccc 649

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA:*

1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/6C.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/6ackfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39.2	6.0	519	2	US-08-249-189-4 Sequence 4, Appli
2	39.2	6.0	519	2	US-08-484-624A-4 Sequence 4, Appli
3	39.2	6.0	519	2	US-08-477-733B-4 Sequence 4, Appli
4	39.2	6.0	519	3	US-09-088-913A-4 Sequence 4, Appli
5	39.2	6.0	519	4	US-08-769-819-4 Sequence 4, Appli
6	39.2	6.0	519	4	US-08-770-974-4 Sequence 4, Appli
7	39.2	6.0	519	4	US-08-770-981-4 Sequence 4, Appli
8	39.2	6.0	519	4	US-09-399-106-4 Sequence 4, Appli
9	39.2	6.0	1004	4	US-09-071-433-85 Sequence 85, Appli
10	39.2	6.0	1004	4	US-09-041-886-26 Sequence 26, Appli
11	39.2	6.0	1004	4	US-08-114-944D-1 Sequence 1, Appli
12	36.4	5.6	459	4	US-09-286-529-7 Sequence 7, Appli
13	36.4	5.6	1164	4	US-08-794-796-1 Sequence 1, Appli
14	36.4	5.6	1347	4	US-09-286-529-18 Sequence 18, Appli
15	36.4	5.6	1859	4	US-09-286-529-19 Sequence 19, Appli
16	35.2	5.4	11282	4	US-09-754-250-3 Sequence 3, Appli
17	35.2	5.4	11282	4	US-08-866-340-29 Sequence 29, Appli
18	35.2	5.4	382	4	US-09-103-875-35 Sequence 35, Appli
19	34.2	5.3	268	1	US-08-039-137-16 Sequence 16, Appli
20	33.8	5.2	933	6	5340934-12 Patent No. 5340934
21	33.3	5.1	1057	3	US-09-188-930-18 Sequence 18, Appli
22	33.3	5.1	1722	4	US-09-142-623-14 Sequence 14, Appli
23	33.3	5.1	2004	3	US-09-188-930-230 Sequence 230, App
24	32.8	5.1	376	4	US-09-149-476-201 Sequence 201, App
25	32.6	5.0	3282	1	US-08-276-852-154 Sequence 154, App
26	32.6	5.0	3282	1	US-08-276-852-169 Sequence 169, App
27	32.6	5.0	3282	1	US-08-899-575-154 Sequence 154, App

C	28	32.6	5.0	3282	1	US-08-899-575-169	Sequence 169, App
C	29	32.6	5.0	3282	1	US-08-899-575-154	Sequence 154, App
C	30	32.6	5.0	3282	1	US-08-899-575-169	Sequence 169, App
C	31	32.6	5.0	3282	5	PCT-US95-08743-154	Sequence 154, App
C	32	32.6	5.0	3282	5	PCT-US95-08743-169	Sequence 169, App
C	33	32.6	5.0	4723	4	US-09-042-353-370	Sequence 370, App
C	34	32.6	5.0	4723	4	US-08-758-417A-218	Sequence 218, App
C	35	32.6	5.0	4926	4	US-09-042-353-418	Sequence 418, App
C	36	32.6	5.0	4926	4	US-08-758-417A-268	Sequence 268, App
C	37	32.6	5.0	5337	2	US-08-784-512-4	Sequence 4, Appli
C	38	32.6	5.0	5337	4	US-09-176-228-4	Sequence 4, Appli
C	39	32.6	5.0	8614	4	US-09-247-352-5	Sequence 5, Appli
C	40	32.6	5.0	8614	4	US-09-466-635-5	Sequence 5, Appli
C	41	32.6	5.0	13254	1	US-08-276-852-156	Sequence 156, App
C	42	32.6	5.0	13254	1	US-08-276-852-170	Sequence 170, App
C	43	32.6	5.0	13254	1	US-08-899-575-156	Sequence 156, App
C	44	32.6	5.0	13254	1	US-08-899-575-170	Sequence 170, App
C	45	32.6	5.0	13254	1	US-08-899-575-156	Sequence 156, App

ALIGNMENTS

RESULT 1
US-08-249-189-4
; Sequence 4, Application US/08249189
; Patent No. 5961974
GENERAL INFORMATION:
APPLICANT: ARMITAGE, RICHARD
APPLICANT: FANSLAW, WILLIAM
APPLICANT: SPRIGGS, MELANIE
APPLICANT: SRINIVASAN, SUBHASHINI
APPLICANT: GIBSON, MARYLOU
TITLE OF INVENTION: NOVEL CYTOKINE
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMUNEX CORPORATION
STREET: 51 UNIVERSITY STREET
CITY: SEATTLE
STATE: WASHINGTON
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple Operating System 7.1
SOFTWARE: Microsoft Word for Apple, version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/249,189
FILING DATE:
CLASSIFICATION: 514
APPLICATION NUMBER: 07/969,703
FILING DATE: October 23, 1992
CLASSIFICATION: 514
APPLICATION NUMBER: 07/805,723
FILING DATE: December 5, 1991
CLASSIFICATION: 514
APPLICATION NUMBER: 07/83,707
FILING DATE: October 25, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia A.
REGISTRATION NUMBER: 34,693
TELECOMMUNICATION INFORMATION:
TELEPHONE: 2065870430
TELEFAX: 2065870606
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 519 base pairs

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAMI0578 row: b column: 10

High quality sequence stop: 673.

FEATURES

Location/Qualifiers

1..700

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4707753"

/clone_lib="NIH_MGC_87"

/tissue_type="mammary adenocarcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: breast; Vector: pCMV-SPORT6; Site:1: NotI;

Site:2: SalI; Cloned unidirectionally; oligo-dr primed.

Average insert size 1.383 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

BASE COUNT 160 a 223 c 189 g 127 t 1 others

ORIGIN

Query Match

57.0%; Score 370; DB 12; Length 700;

Best Local Similarity 99.5%;

Pred. No. 2e-82;

Mismatches 0; Indels 2; Gaps 2;

Matches 392; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

```
QY 257 GGATCGCTGCTCAGAGGCGTCCGAGGAGAGGACCAATCCTGTAGTGAAGCTGG 316
    |||||||
DB 84 GGATCGCTGCTCAGAGGCGTCCGAGGAGAGGACCAATCCTGTAGTGAAGCTGG 316
    |||||||
QY 317 GAGCCTCGAAGGCCCATCTACTTCCCTGACTTGTGTACAGCCACTGCTACCCATTCT 376
    |||||||
DB 144 GAGCCTCGAAGGCCCATCTACTTCCCTGACTTGTGTACAGCCACTGCTACCCATTCT 376
    |||||||
QY 377 GGAGATGTTTCCCGAGTATCCAGTGGGCTCCCGGAGGCCCATGTTTGAGGAGGGGTG 436
    |||||||
DB 204 GGAGATGTTTCCCGAGTATCCAGTGGGCTCCCGGAGGCCCATGTTTGAGGAGGGGTG 436
    |||||||
QY 437 CCGCAACAGCAGAGTCTCTGACCTGACCAAGGAGCCGAGTTGGAACCGGGGAGCAG 496
    |||||||
DB 264 CCGCAACAGCAGAGTCTCTGACCTGACCAAGGAGCCGAGTTGGAACCGGGGAGCAG 496
    |||||||
QY 497 AGCCAGGTGGCCACGATGACCAATGCAATGCAACCGGGGCTATGACTATCACT 556
    |||||||
DB 324 AGCCAGGTGGCCACGATGACCAATGCAATGCAACCGGGGCTATGACTATCACT 556
    |||||||
QY 557 GGCAGC-ATCTACATCTACATGAGCAGTACGAGGAGGAGCAGGAGGAGGAGGAGGAG 615
    |||||||
DB 384 GGCAGC-ATCTACATCTACATGAGCAGTACGAGGAGGAGGAGGAGGAGGAGGAGGAG 615
    |||||||
QY 616 CCCAGCTACCCCGAGACCTCCATACCCCATTTCCC 649
    |||||||
DB 444 CCCAGCTA-CCCGAGACCTCCATACCCCATTTCCC 476
    |||||||
```

Search completed: April 15, 2003, 16:26:25
Job time : 756.63 secs

/note="Organ: ovary; Vector: pOTB7; Site.1: XhoI; Site.2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 181 a 297 c 267 g 156 t

ORIGIN

Query Match 58.7%; Score 381; DB 12; Length 901;
Best Local Similarity 99.7%; Pred. No. 3.9e-85;
Matches 392; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

257 GGATCGCTCTCCAGAGGCGCTCCGAGGAGAGGACCCAACTCTGTAGCTGGAGCTGG 316
|||||
55 GGATCGCTCTCCAGAGGCGGT-CGCGAGGAGAGGAGGACCAATCTGTAGCTGGAGCTGG 113
|||||
317 GAGCCTCGGAGGCGCCATCATCTCCGCTGACTGTACAGCCACTGTACCATTTCT 376
|||||
114 GAGCCTCGGAGGCGCCATCATCTCCGCTGACTGTACAGCCACTGTACCATTTCT 173
|||||
377 GGAGATGTTTCCCACTATCCACTGGGCTCCCGAGCCCAAGTTTGGAGCGAGGGTG 436
|||||
174 GGAGATGTTTCCCACTATCCACTGGGCTCCCGAGCCCAAGTTTGGAGCGAGGGTG 233
|||||
437 CGGCAACAGCAGAGCTCTGTGACGTACAGGAGGAGCCGAGTGTGAACCCGGGAGCAG 496
|||||
234 CGGCAACAGCAGAGCTCTGTGACGTACAGGAGGAGCCGAGTGTGAACCCGGGAGCAG 293
|||||
497 AGCCAGGTGGCCAGGCTACCATGTGATGATGATGATGATGATGATGATGATGATGAT 556
|||||
294 AGCCAGGTGGCCAGGCTACCATGTGATGATGATGATGATGATGATGATGATGATGAT 353
|||||
557 GGCACATCTACATCTACATGAGCAGGAGTGGGGGAGCACCAGGCTCTGAGAGACTC 616
|||||
354 GGCACATCTACATCTACATGAGCAGGAGTGGGGGAGCACCAGGCTCTGAGAGACTC 413
|||||
617 CCAGCTACCCCGAACCCTCATACCCCATTTCCC 649
|||||
414 CCAGCTACCCCGAACCCTCATACCCCATTTCCC 446
|||||

RESULT 14 954 bp mRNA linear EST 20-AUG-2002
B0930450
LOCUS B0930450
DEFINITION AGENCOURT_8946139 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6462494
5', mRNA sequence.
ACCESSION B0930450
VERSION B0930450.1 GI:22345481
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 954)
AUTHORS NIH-MGC http://mhc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2652 row: h column: 15
High quality sequence stop: 576.
Location/Qualifiers
1. 954

FEATURES
source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6462494"
/clone_11b="NIH_MGC_101"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site.1: EcoRI; Site.2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."

BASE COUNT 205 a 320 c 278 g 151 t

ORIGIN

Query Match 58.6%; Score 380.4; DB 14; Length 954;
Best Local Similarity 99.7%; Pred. No. 5.6e-85;
Matches 381; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

268 CAAGAGCGCTCCGAGGAGAGGAGGAGCCAACTCTGTAGCTGGAAGCTGGAGCTCGGA 327
|||||
1 CAAGAGCGCTCCGAGGAGAGGAGGAGCCAACTCTGTAGCTGGAAGCTGGAGCTCGGA 60
|||||
328 GGGCCATCCATCTCTCCGACTGTGTACAGCCACTGTACCATTTCTGAGATGTTTC 387
|||||
61 GGGCCATCCATCTCTCCGACTGTGTACAGCCACTGTACCATTTCTGAGATGTTTC 120
|||||
388 CCCAGTATCACTGGGCTCCCGAGCCCAAGTTTGAAGGAGGGGTCCCGAACAGA 447
|||||
121 CCAGTATCACTGGGCTCCCGAGCCCAAGTTTGAAGGAGGGGTCCCGAACAGA 180
|||||
448 GAGTCTCTGAGCCTTACAGGAGGAGCGAGTTGGAACCCGGGAGCAGAGCGAGTGGC 507
|||||
181 GAGTCTCTGAGCCTTACAGGAGGAGCGAGTTGGAACCCGGGAGCAGAGCGAGTGGC 240
|||||
508 CCAGGTTACCAATGGCATTCATGTACACCGGCGGTCTATGACTATTCATTCAGCAATCTA 567
|||||
241 CCAGGTTACCAATGGCATTCATGTACACCGGCGGTCTATGACTATTCAGCAATCTA 300
|||||
568 CATCTCAATATGACACAGTACTGGGGGAGCACCAGGCTCTGAGAGCTCCAGCTACCC 627
|||||
301 CATCTCAATATGACACAGTACTGGGGGAGCACCAGGCTCTGAGAGCTCCAGCTACCC 360
|||||
628 CGAACCTCCATACCCCATTTCCC 649
|||||
361 CGAACCTCCATACCCCATTTCCC 382
|||||

RESULT 15 700 bp mRNA linear EST 10-APR-2001
BG576875
LOCUS BG576875
DEFINITION 60259233F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:4707753 5',
mRNA sequence.
ACCESSION BG576875
VERSION BG576875.1 GI:13584528
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 700)
AUTHORS NIH-MGC http://mhc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DP
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be

QY 437 CCGCAACACAGAGTCTCTGACCTGACACAGGAGCGCAGTTGGACCCGGGGAGCAG 496
 |||||||
 DB 194 CCGCAACACAGAGTCTCTGACCTGACACAGGAGCGCAGTTGGACCCGGGGAGCAG 253
 |||||||
 QY 497 ACCCAGGTGGCCAGGATACCAATGGCATTCATGTCACCGGGGCTATGATGATACAT 556
 |||||||
 DB 254 ACCCAGGTGGCCAGGATACCAATGGCATTCATGTCACCGGGGCTATGATGATACAT 313
 |||||||
 QY 557 GGCACATCTACATCTACATGACACAGTACGTGGGGGACACCGGGTCTTGAGACCTC 616
 |||||||
 DB 314 GGCACATCTACATCTACATGACACAGTACGTGGGGGACACCGGGTCTTGAGACCTC 373
 |||||||
 QY 617 CCAGCTACCCCGAACCTCCATACCCCATTTCCC 649
 |||||||
 DB 374 CCAGCTACCCCGAACCTCCATACCCCATTTCCC 406
 |||||||

RESULT 10
 AM630662
 LOCUS
 DEFINITION
 hbs8h07.y1 NCI-CGAP-GUI Homo sapiens CDNA clone IMAGE:2969629 5' similar to gb:L04270 TUMOR NECROSIS FACTOR RECEPTOR 2 RELATED PROTEIN PRECURSOR (HUMAN); contains element MSRI repetitive element ;, mRNA sequence.

ACCESSION
 AM630662
 VERSION
 AM630662.1 GI:7377452
 EST.

KEYWORDS
 EST.

SOURCE
 HOMO
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
 1 (bases 1 to 427)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 Unpublished (1997)

JOURNAL
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Chris Moskalko, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: Christa Prange, The I.M.A.G.E. Consortium DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov/image/html/resources.shtml
 Seq primer: -40RP from Gibco
 High quality sequence stop: 340.

FEATURES
 source
 1..427
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2969629"
 /clone_lib="NCI-CGAP-GUI"
 /tissue_type="2 pooled high-grade transitional cell tumors"
 /lab_host="DH10B"
 /note="Organ: genitourinary tract; Vector: pCMV-SPORT6; Site_1: Salt; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies."

BASE COUNT
 89 a 145 c 118 g 75 t

ORIGIN
 Query Match 59.8%; Score 388; DB 10; Length 427;
 Best Local Similarity 100.0%; Pred. No. 5e-87;
 Matches 388; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 262 GGTGCTCAAGAGCGCTCCGACGAGAGGAGGACCAATCTCTGATGCTGGAAGCTGGAGCC 321
 |||||||
 DB 1 GGTGCTCAAGAGCGCTCCGACGAGAGGAGGACCAATCTCTGATGCTGGAAGCTGGAGCC 60
 |||||||

QY 322 TCCGAAGGCCCATTCATCTCCCTGACTTGGTACAGCCACTGACCCATTTCTGAGA 381
 |||||||
 DB 61 TCCGAAGGCCCATTCATCTCCCTGACTTGGTACAGCCACTGACCCATTTCTGAGA 120
 |||||||
 QY 382 TGTTCCTCCAGATATCCACTGAGGCTCCCGCAGCCAGTTTGGAGGAGGGGTGCCGA 441
 |||||||
 DB 121 TGTTCCTCCAGATATCCACTGAGGCTCCCGCAGCCAGTTTGGAGGAGGGGTGCCGA 180
 |||||||
 QY 442 ACAGCAGAGTCTCTTGAGACTGACACGAGGAGCCGACGTTGGAAACCCGGGAGCAGAGCCA 501
 |||||||
 DB 181 ACAGCAGAGTCTCTTGAGACTGACACGAGGAGCCGACGTTGGAAACCCGGGAGCAGAGCCA 240
 |||||||
 QY 502 GGTGGCCAGGAGTACCAATGAGCATTCATGTCACCGGGGATGATGATGATGATGATGATGAT 561
 |||||||
 DB 241 GGTGGCCAGGAGTACCAATGAGCATTCATGTCACCGGGGATGATGATGATGATGATGATGAT 300
 |||||||
 QY 562 CATCTACATCTACATGACACAGTACTGGGGGACACCGGGTCTTGAGACCTCCAGC 621
 |||||||
 DB 301 CATCTACATCTACATGACACAGTACTGGGGGACACCGGGTCTTGAGACCTCCAGC 360
 |||||||

QY 622 TACCCCGAAGCTCCATACCCCATTTCCC 649
 |||||||
 DB 361 TACCCCGAAGCTCCATACCCCATTTCCC 388
 |||||||

RESULT 11
 BG491235
 LOCUS
 DEFINITION
 602535283F1 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:4654076 5', mRNA sequence.

ACCESSION
 BG491235
 VERSION
 BG491235.1 GI:13452747
 EST.

KEYWORDS
 EST.

SOURCE
 HOMO
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
 1 (bases 1 to 640)
 NIH-MGC <http://mhc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)

JOURNAL
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCPD/DMP
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: L1CM1441 row: e column: 21
 High quality sequence start: 199
 High quality sequence stop: 640.

FEATURES
 source
 1..640
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4654076"
 /clone_lib="NIH_MGC_41"
 /tissue_type="amelanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."

BASE COUNT
 134 a 205 c 159 g 142 t

ORIGIN
 Query Match 58.7%; Score 381; DB 12; Length 640;

```

Db 432 GGCACATCTACATCTACATGAGCCAGTACTGGGGGACACCGGCTCTGAGACTC 491
QY 617 CCAGCTACCCCGACACTCATACCCATTCCC 649
Db 492 CCAGCTACCCCGACACTCATACCCATTCCC 524

RESULT 8
LOCUS BG036632 920 bp mRNA linear EST 24-JAN-2001
DEFINITION 60236834F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:442825 5',
mRNA sequence.
ACCESSION BG036632
VERSION BG036632.1 GI:12432013
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 920)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM0178 row: c column: 10
High quality sequence stop: 695.
Location/Qualifiers
1. 920
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:442825"
/tissue="NIH_MGC_91"
/issue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.4 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 187 a 304 c 270 g 159 t
ORIGIN
Query Match 60.6%; Score 393; DB 12; Length 920;
Best Local Similarity 100.0%; Pred. No. 3.8e-88;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 GGATCGCTGCTCAAGAGCGCTCCGAGGAGAGGAGCCCAATCTGTAGTGAAGCTGG 316
Db 52 GGATCGCTGCTCAAGAGCGCTCCGAGGAGAGGAGCCCAATCTGTAGTGAAGCTGG 111
QY 317 GAGCTCCGGAAGGCCCATCTACTCTCTGACTTGGTGAAGCCACTGCTACCATTTCT 376
Db 112 GAGCTCCGGAAGGCCCATCTACTCTCTGACTTGGTGAAGCCACTGCTACCATTTCT 171
QY 377 GGAATGTTTCCAGATCCAGTACAGTGGGCTCCGAGGCCAGCTTTGGAGGAGGGTG 436
Db 172 GGAATGTTTCCAGATCCAGTACAGTGGGCTCCGAGGCCAGCTTTGGAGGAGGGTG 231
QY 437 CCGCAACAGCAGAGTCTCTGACCTGACACGAGGAGCCAGTGTGAACCCGGGGAGCAG 496
Db 232 CCGCAACAGCAGAGTCTCTGACCTGACACGAGGAGCCAGTGTGAACCCGGGGAGCAG 291
QY 497 AGCCAGGTGGCCACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 556

```

```

Db 292 AGCCAGTGGCCCGCCAGCTACCATGATGATGATGATGATGATGATGATGATGAT 351
QY 557 GGCACATCTACATCTACATGAGCCAGTACTGGGGGACACCGGCTCTGAGACTC 616
Db 352 GGCACATCTACATCTACATGAGCCAGTACTGGGGGACACCGGCTCTGAGACTC 411
QY 617 CCAGCTACCCCGACACTCATACCCATTCCC 649
Db 412 CCAGCTACCCCGACACTCATACCCATTCCC 444

RESULT 9
LOCUS B1821009 954 bp mRNA linear EST 04-OCT-2001
DEFINITION 603035664F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5176935 5',
mRNA sequence.
ACCESSION B1821009
VERSION B1821009.1 GI:15932559
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 954)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM1440 row: o column: 16
High quality sequence stop: 758.
Location/Qualifiers
1. 954
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5176935"
/tissue="NIH_MGC_115"
/lab_host="DH10B"
/notes="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

BASE COUNT 212 a 313 c 281 g 148 t
ORIGIN
Query Match 60.6%; Score 393; DB 13; Length 954;
Best Local Similarity 100.0%; Pred. No. 3.9e-88;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 GGATCGCTGCTCAAGAGCGCTCCGAGGAGAGGAGCCCAATCTGTAGTGAAGCTGG 316
Db 14 GGATCGCTGCTCAAGAGCGCTCCGAGGAGAGGAGCCCAATCTGTAGTGAAGCTGG 73
QY 317 GAGCTCCGGAAGGCCCATCTACTCTCTGACTTGGTGAAGCCACTGCTACCATTTCT 376
Db 74 GAGCTCCGGAAGGCCCATCTACTCTCTGACTTGGTGAAGCCACTGCTACCATTTCT 133
QY 377 GGAATGTTTCCAGATCCAGTACAGTGGGCTCCGAGGCCAGCTTTGGAGGAGGGTG 436
Db 134 GGAATGTTTCCAGATCCAGTACAGTGGGCTCCGAGGCCAGCTTTGGAGGAGGGTG 193

```

DEFINITION 601122062F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3346311 5',
mRNA sequence.
ACCESSION BE275240
VERSION BE275240.1 GI:9150193
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 868)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L10M135 row: K column: 16
High quality sequence stop: 746.
Location/Qualifiers
1..868
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3346311"
/clone_lib="NIH_MGC_20"
/tissue_type="melanocytic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene, Berkeley) and Superscript II RT (Life Technologies)."
BASE COUNT 170 a 286 c 253 g 159 t
ORIGIN
Query Match 60.6%; Score 393; DB 10; Length 868;
Best Local Similarity 100.0%; Pred. No. 3.7e-88;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 7
B0957322 895 bp mRNA linear EST 21-AUG-2002
LOCUS
DEFINITION AGC0977754 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6384424
ACCESSION B0957322
VERSION B0957322.1 GI:22372800
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 895)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L10M2579 row: g column: 17
High quality sequence stop: 645.
Location/Qualifiers
1..895
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6384424"
/clone_lib="NIH_MGC_40"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: prostate; Vector: pORF7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
Note: this is a NIH_MGC Library."
BASE COUNT 190 a 291 c 252 g 159 t 3 others
ORIGIN
Query Match 60.6%; Score 393; DB 14; Length 895;
Best Local Similarity 100.0%; Pred. No. 3.8e-88;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsungemail.kr@kribb.re.kr
 Plate: 16 row: E column: 09
 High quality sequence stop: 538.

FEATURES

source
 1. 538
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="S12SN0216-16-E09"
 /clone_id="S12SN0216"
 /sex="F"
 /tissue_type="Lymph node"
 /cell_type="Epithelial"
 /cell_line="SNU-216"
 /lab_host="Top10F"
 /note="Organ: Stomach; Vector: PCNS; Site_1: EcoRI; Site_2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transfection of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

BASE COUNT 116 a 185 c 144 g 93 t

ORIGIN

Query Match 64.4%; Score 418; DB 14; Length 538;
 Best Local Similarity 100.0%; Pred. No. 1.6e-94;
 Matches 418; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGAGTGTACACATCGGAGCTACTTTCGTGACGCGCGCTGGCACTGAAGCCGAGCTAA 60
 DB 121 CGAGTGTACACATCGGAGCTACTTTCGTGACGCGCGCTGGCACTGAAGCCGAGCTAA 180
 QY 61 AGATGAAGTTGGAAAGGTTAACAACCACTGCGTCCCTGCAAGGAGGAGGCACTTCAGAA 120
 DB 181 AGATGAAGTTGGAAAGGTTAACAACCACTGCGTCCCTGCAAGGAGGAGGCACTTCAGAA 240
 QY 121 TACCTCTCCCGCAGCGCGCTGCGACGCCACACAGGTGTGAAACCAAGGTCTGGT 180
 DB 241 TACCTCTCCCGCAGCGCGCTGCGACGCCACACAGGTGTGAAACCAAGGTCTGGT 300
 QY 181 GGAGGACACTCCAGGCACTGGCCAGTCCGACAAACCGCAAAATTCATAGAGCCACT 240
 DB 301 GGAGGACACTCCAGGCACTGGCCAGTCCGACAAACCGCAAAATTCATAGAGCCACT 360
 QY 241 GCCCCAGAGATGTCAGAGATGCTGCTCAAGAGGCTCCGAGGAGGAGAGACCAATCC 300
 DB 361 GCCCCAGAGATGTCAGAGATGCTGCTCAAGAGGCTCCGAGGAGGAGAGACCAATCC 420
 QY 301 TGTAGCTGGAAGCTGGAGGCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
 DB 421 TGTAGCTGGAAGCTGGAGGCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
 QY 361 ACTGCTACCACTTCTGAGATGTTCCCGAGTATCCATGGGCGCCCGAGAGCCCA 418
 DB 481 ACTGCTACCACTTCTGAGATGTTCCCGAGTATCCATGGGCGCCCGAGAGCCCA 538

RESULT 5
 AUI18203 814 bp mRNA linear EST 01-AUG-2002
 LOCUS AUI18203
 DEFINITION AUI18203 HEMBA1 Homo sapiens cDNA clone HEMBA1003089 5', mRNA

sequence.
 ACCESSION AUI18203
 VERSION AUI18203.1 GI:10933231
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE

1 (bases 1 to 814)
 Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and

Isogai,T.
 HRI human cDNA project
 Unpublished (2000)
 Contact: Takao Isogai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: genom@hri.co.jp

HRI human cDNA project: 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.
 location/Qualifiers
 1. 814

FEATURES

source
 1. 814
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HEMBA1003089"
 /clone_id="HEMBA1"
 /tissue_type="whole embryo, mainly head"
 /dev_stage="embryo, 10 weeks"
 /note="Vector: pME18SFL3"

BASE COUNT 173 a 255 c 232 g 151 t 3 others

ORIGIN

Query Match 60.6%; Score 393; DB 9; Length 814;
 Best Local Similarity 100.0%; Pred. No. 3.6e-88;
 Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 GGATCGCTGCTCAAGAGGCGTCCGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 316
 DB 98 GGATCGCTGCTCAAGAGGCGTCCGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 157
 QY 317 GAGCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 376
 DB 158 GAGCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 217
 QY 377 GGAGATGTTTCCCGAGTATCCAGTATCCAGTATCCAGTATCCAGTATCCAGTATCCAGTATCC 436
 DB 218 GGAGATGTTTCCCGAGTATCCAGTATCCAGTATCCAGTATCCAGTATCCAGTATCCAGTATCC 277
 QY 437 CCGCAACAGCAGAGTCTCTGTGACCTGACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 496
 DB 278 CCGCAACAGCAGAGTCTCTGTGACCTGACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 337
 QY 497 AGCCAGGTGGGCCACGAGTATCCAGTATCCAGTATCCAGTATCCAGTATCCAGTATCCAGTATCC 556
 DB 338 AGCCAGGTGGGCCACGAGTATCCAGTATCCAGTATCCAGTATCCAGTATCCAGTATCCAGTATCC 397
 QY 557 GGCAACATCTACATCTACATCTACATCTACATCTACATCTACATCTACATCTACATCTACATCT 616
 DB 398 GGCAACATCTACATCTACATCTACATCTACATCTACATCTACATCTACATCTACATCTACATCT 457
 QY 617 CCAGCTACCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 649
 DB 458 CCAGCTACCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 490

RESULT 6
 BE275240 868 bp mRNA linear EST 13-JUL-2000
 LOCUS BE275240

OY 292 ACCCAATCTGTAGCTGAGAGCTGGAGCTCCGAAAGGCCATCTACTTCCCTGACT 351
 DB 301 ACCCAATCTGTAGCTGAGAGCTGGAGCTCCGAAAGGCCATCTACTTCCCTGACT 360
 OY 352 GGTACAGCCAGCTGCTACCTATTTCTGGAGATGTTTCCAGATCCACTGGGCTCCCGC 411
 DB 361 GGTACAGCCAGCTGCTACCTATTTCTGGAGATGTTTCCAGATCCACTGGGCTCCCGC 420
 OY 412 AGCCCAAGTTTGGAGAGCGAGGCTGCCGACAGCAGAGTCTCTGGACCTTACAGGGA 471
 DB 421 AGCCCAAGTTTGGAGAGCGAGGCTGCCGACAGCAGAGTCTCTGGACCTTACAGGGA 480
 OY 472 GCGCAGTTTGAACCCCGGAGAGCAGAGCAGTGGCCACAGTACCAATGGCATTTATGT 531
 DB 481 GCGCAGTTTGAACCCCGGAGAGCAGAGCAGTGGCCACAGTACCAATGGCATTTATGT 540
 OY 532 CACCGCGGGTCTATGATCATCTGCAACATCTACATCTACAAATGGACAGTACTGGG 591
 DB 541 CACCGCGGGTCTATGATCATCTGCAACATCTACATCTACAAATGGACAGTACTGGG 600
 OY 592 GGGACACCGGGTCTGAGAGACCTCCAGTACCCCGAAGCTCCATACCCCATTTCC 649
 DB 601 GGGACACCGGGTCTGAGAGACCTCCAGTACCCCGAAGCTCCATACCCCATTTCC 658
 RESULT 3
 BI769498 900 bp mRNA linear EST 25-SEP-2001
 LOCUS 603059047.1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5208394 5',
 DEFINITION mRNA sequence.
 ACCESSION BI769498
 VERSION BI769498.1 GI:15761076
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 900)
 NIH-MGC http://mgc.ncl.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgaabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM11522 row: n column: 11
 High quality sequence stop: 834.
 Location/Qualifiers
 . 900
 1. 900
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5208394"
 /clone_id="NIH_MGC_122"
 /lab_host="DH10B"
 /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
 Site:1: NotI; Site:2: EcoRV (destroyed); RNA source
 anonymous pool of 24 week female lung, 16 week female
 spleen, and 20-22 week male spleens. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.4 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 026. Note:
 this is a NIH-MGC Library."
 BASE COUNT 199 a 298 c 246 g 157 t
 ORIGIN
 Query Match 64.9%; Score 421.4; DB 13; Length 900;

Best Local Similarity 83.4%; Pred. No. 2 8e-95;
 Matches 551; Conservative 0; Mismatches 1; Indels 109; Gaps 2;
 OY 97 CTGCAAGCGCAGGCGACTTCCAGAAATACCTCTCTCCCGAGCGCCGCTGCCAGCCACAC 156
 DB 1. CTGCAAGCGCAGGCGACTTCCAGAAATACCTCTCTCCCGAGCGCCGCTGCCAGCCACAC 60
 OY 157 CAGGTGTGAGAACCAAGTCTGTGGAGAGCAGCTCCAGGCACTGGCCAGTCCGACACAC 216
 DB 61 CAGGTGTGAGAACCAAGTCTGTGGAGAGCAGCTCCAGGCACTGGCCAGTCCGACACAC 120
 OY 217 CTGCAAAATTCATAGAGCCACTGCCCCAGAGATGTCA----- 256
 DB 121 CTGCAAAATTCATAGAGCCACTGCCCCAGAGATGTCAAGAACCATGTGATGTGGC 180
 OY 257 ----- 256
 DB 181 CGTCTGTGCGCAGCTGCGCTTCTTGTGCTCTTCCACCGTCTTCTCTGCAATCTGGA 240
 OY 257 -----GGATCGCTGCTCAAGAGCGCTCCGAGAGGA 288
 DB 241 GAGCCACCTTCTCTCTGCGAGAAACTGGATCGCTGCTCAAGAGCGCTCCGAGAGGA 300
 OY 289 GGGACCCATCTCTGTAGCTGAGAGCTGGAGGCTCCGAAAGGCCATCTACTTCCCTGA 348
 DB 301 GGGACCCATCTCTGTAGCTGAGAGCTGGAGGCTCCGAAAGGCCATCTACTTCCCTGA 360
 OY 349 CTGTGTGAGAGCCAGCTGCTACCTATTTCTGGAGATGTTTCCCGATTCACACTGGGCTCC 408
 DB 361 CTGTGTGAGAGCCAGCTGCTACCTATTTCTGGAGATGTTTCCCGATTCACACTGGGCTCC 420
 OY 409 CCGAGCCCGAGTTTGGAGAGCGAGGCTGCCGACAGCAGAGTCTCTGGACCTTACAGG 468
 DB 421 CCGAGCCCGAGTTTGGAGAGCGAGGCTGCCGACAGCAGAGTCTCTGGACCTTACAGG 479
 OY 469 GGAAGCCGAGTTTGGAGAGCCCGGAGAGCAGAGGCTGGCCACGTTACCAATGGGATTA 528
 DB 480 GGAAGCCGAGTTTGGAGAGCCCGGAGAGCAGAGGCTGGCCACGTTACCAATGGGATTA 539
 OY 529 TGTACACCGCGGGTCTATGATCATCTACTGCAACATCTCTACATCTCAATGGACAGTACT 588
 DB 540 TGTACACCGCGGGTCTATGATCATCTACTGCAACATCTCTACATCTCAATGGACAGTACT 599
 OY 589 GGGGGAGCACCAGCGGTCTCTGAGAGCCTCCAGTACCCCGAAGCTCCATACCCCATTTCC 648
 DB 600 GGGGGAGCACCAGCGGTCTCTGAGAGCCTCCAGTACCCCGAAGCTCCATACCCCATTTCC 659
 OY 649 C 649
 DB 660 C 660
 RESULT 4
 BM764302 538 bp mRNA linear EST 04-MAR-2002
 LOCUS K-EST0045797 S125NU216 Homo sapiens cDNA clone S125NU216-16-E09 5',
 DEFINITION mRNA sequence.
 ACCESSION BM764302
 VERSION BM764302.1 GI:19093917
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 538)
 Kim, N.S., Hahn, Y., Oh, J. H., Lee, J. Y., Ahn, H. Y., Chu, M. Y., Kim, M. R.,
 Oh, K. J., Cheong, J. E., Sohn, H. Y., Kim, J. M., Park, H. S., Kim, S. and
 Kim, Y. S.
 21C Frontier Korean EST Project 2001
 TITLE Unpublished (2002)
 JOURNAL Contact: Kim YS
 COMMENT Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology


```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:518684"
/clone_lib="NIH_MGC_116"
/lab_host="DH10B"
/Note="Organ: pooled colon, kidney, stomach; Vector:
PCMV-SpOrf6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb.
Insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH MGC Library."

BASE COUNT      156 a      233 c      198 g      141 t
ORIGIN
Query Match      71.2%; Score 462.4; DB 13; Length 728;
Best Local Similarity 84.3%; Pred. No. 1.3e-105;
Matches 592; Conservative 0; Mismatches 1; Indels 109; Gaps 2;

QY 1 CGAGGTACACACTGCGAGCTACTTCTGACTGCCGCCCTGCGACTGAGCGGACTCAA 60
    |||||||
Db 28 CGAGGTACACACTGCGAGCTACTTCTGACTGCCGCCCTGCGACTGAGCGGACTCAA 87

QY 61 AGATGAAAGTTGGGAGGGTAAACAACCAACGTCGTCCTCGAAGGAGGCGACTTCCAGAA 120
    |||||||
Db 88 AGATGAAAGTTGGGAGGGTAAACAACCAACGTCGTCCTCGAAGGAGGCGACTTCCAGAA 147

QY 121 TACCTCCTCCCGAGGCCGCTGCCAGCCCAACCAAGGTGTAGAACCAAGGTCGTGT 180
    |||||||
Db 148 TACCTCCTCCCGAGGCCGCTGCCAGCCCAACCAAGGTGTAGAACCAAGGTCGTGT 207

QY 181 GGAGGACAGCTCCAGGACATGCCATGCCAGACAACTGCAAAAATTCATTAGAGCCACT 240
    |||||||
Db 208 GGAGGACAGCTCCAGGACATGCCATGCCAGACAACTGCAAAAATTCATTAGAGCCACT 267

QY 241 GCGCCCGAGATGTCA----- 256
    |||||||
Db 268 GCGCCCGAGATGTCAAGAACCAATGCTGATGTGGCCGCTTCTGCGCCACTGGCCCTTCTT 327

QY 257 ----- 256
    |||||||
Db 328 TCTGCTCCTTGGCAGCGTCTTCTGCTGCATCTGGAAGAGCCACCTTCTCTGCGAGAA 387

QY 257 -----GATCGCTGCTCAAGAGCGCTCCGAGGAGAGGAGCAATCTGTAGCTGGAAG 312
    |||||||
Db 388 ACTGGGATCGCTGCTCAAGAGCGCTCCGAGGAGAGGAGCAATCTGTAGCTGGAAG 447

QY 313 CTGGGAGGCTCCGAAAGGCCATCATTAATCTCCGACTTGTGACAGCAATGCTTACCCAT 372
    |||||||
Db 448 CTGGGAGGCTCCGAAAGGCCATCATTAATCTCCGACTTGTGACAGCAATGCTTACCCAT 507

QY 373 TTCTGGAGATTTTCCCACTATCACTGGGCTCCCGAGCCCAAGTTTGGAGGAGG 432
    |||||||
Db 508 TTCTGGAGATTTTCCCACTATCACTGGGCTCCCGAGCCCAAGTTTGGAGGAGG 567

QY 433 GGTGCGGCAACAGACAGATCTCTGACCTGACGAGGAGCCGAGTTGGAAACCCGGGGA 492
    |||||||
Db 568 GGTGCGGCAACAGACAGATCTCTGACCTGACGAGGAGCCGAGTTGGAAACCCGGGGA 626

QY 493 GCAGAGCAGAGTGGCCAGCGTACCAATGATTCATGTACCGGCGGCTGTATGACTAT 552
    |||||||
Db 627 GCAGAGCAGAGTGGCCAGCGTACCAATGATTCATGTACCGGCGGCTGTATGACTAT 686

QY 553 CACTGCAACATCTACATCTACAAATGAGACGTAAGTGGGGG 594
    |||||||
Db 687 CACTGCAACATCTACATCTACAAATGAGACGTAAGTGGGGG 728

RESULT 2

```

```

B0898015
LOCUS      B0898015      968 bp      mRNA      linear      EST 16-AUG-2002
DEFINITION AGENCOURT_8125942 Lupski_dorsal_root_ganglion Homo sapiens cDNA
clone IMAGE:6177574 5', mRNA sequence.
ACCESSION  B0898015
VERSION    B0898015.1 GI:22290029
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  1 (bases 1 to 968)
            NIH-MGC http://mhc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            JOURNAL Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapds-remail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
            cDNA Library Preparation: Life Technologies, Inc.
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
            DNA sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNLN at:
            http://image.llnl.gov
            Plate: L1AM1355 row: 1 column: 23
            High quality sequence stop: 615.

FEATURES
    source
        location/Qualifiers
            1..968
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:6177574"
                /clone_lib="Lupski_dorsal_root_ganglion"
                /sex="male"
                /tissue_type="dorsal root ganglia"
                /dev_stage="adult, 36 yr"
                /lab_host="DH10B"
                /note="Vector: PCMV-SpOrf6 (Life Technologies); Site_1:
                NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                Directionally cloned using the following adaptors:
                5'-TCGACCCAGCGCTCCG-3' and
                5'-GACTACTTCTAGATCCGACGCGCCGCTT(15)-3'. Size selected >
                1 kb for average insert length 1.7 kb. This is a primary
                library, non-amplified. Library constructed by Life
                Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
                College of Medicine) and is available through Life
                Technologies."
BASE COUNT      219 a      320 c      260 g      168 t      1 others
ORIGIN
Query Match      66.3%; Score 430.4; DB 14; Length 968;
Best Local Similarity 83.4%; Pred. No. 1.6e-97;
Matches 549; Conservative 0; Mismatches 1; Indels 108; Gaps 1;

QY 100 CAAGGAGGCGACTTCCAGATACCTCTCCGCCAGCGCCGCTGCCAGGCCACACAG 159
    |||||||
Db 1 CAAGGCGGCGACTTCCAGATACCTCTCCGCCAGCGCCGCTGCCAGGCCACACAG 60

QY 160 GTGTGAACCAAGTGTGTGTGAGGAGCTCCAGGACGCTCCGACATGTCGACAACTG 219
    |||||||
Db 61 GTGTGAACCAAGTGTGTGTGAGGAGCTCCAGGACGCTCCGACATGTCGACAACTG 120

QY 220 CAAATAATCAATAGAGCACTGCCCCAGAGATGTCA----- 256
    |||||||
Db 121 CAAATAATCAATAGAGCACTGCCCCAGAGATGTCAAGAACCATGTGTATGTCGCGGT 180

QY 257 ----- 256
    |||||||
Db 181 TCTGTCGACACTGCGCTTCTTCTGCTCTTCCACCGCTTCTCTGCAATCTGGAAGAG 240

QY 257 -----GGATGCGTGTCAAGAGCGTCCGAGGAGAGGAGG 291
    |||||||
Db 241 CCACCTTCTCTGCGAGAAACTGGGATGCTGCTCAAGAGGCGTCCGAGGAGAGGAGG 300

```

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:09:45 ; Search time 749.63 Seconds
(without alignments)
14021.414 Million cell updates/sec

Title: US-09-917-372-2_COPY_629_1277

Perfect score: 649

Sequence: 1 cgaagtacacactgcagac.....aacctccatcccatcc

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 segs, 8097743376 residues 32308132

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Database:

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

EST:*
1: em_estba:*
2: em_esthm:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrl:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	462.4	71.2	728	13	BT65703 603046670
2	430.4	66.3	968	14	B0898015 AGENCOURT
3	421.4	64.9	900	13	BT69498 603059047
4	418	64.4	538	14	BM764302 K-EST0045
5	393	60.6	814	9	AU118203
6	393	60.6	868	10	BE275240 601122062

7	393	60.6	895	14	B0957322
8	393	60.6	920	12	B0366332
9	393	60.6	954	13	B1821009
10	388	59.8	427	10	AM630662
11	381	58.7	640	12	BG491235
12	381	58.7	714	13	B1829057
13	381	58.7	901	12	BE740143
14	380.4	58.6	954	14	B0930450
15	370	57.0	700	12	BG576875
16	355	54.7	749	12	BE789430
17	347	53.5	772	12	BE765684
18	346.6	53.0	579	10	BE207591
19	344	53.0	735	12	BG331666
20	343.8	53.0	606	13	B1905922
21	342	52.7	598	12	BG403820
22	322	49.6	801	12	BF792867
23	316	48.7	458	14	BM766807
24	316	48.7	501	14	BM766833
25	316	48.7	537	14	BM766352
26	316	48.7	572	14	BM767287
27	316	48.7	666	14	BM767318
28	314.4	48.4	634	14	BM746911
29	311	47.9	598	14	BM723050
30	301.8	46.5	360	14	BM766382
31	293.4	45.2	386	9	AA099514
32	282	43.5	598	14	B0305454
33	275	42.4	506	14	BM855039
34	274.4	42.3	276	14	BM784961
35	274	42.2	808	13	B1818411
36	273.4	42.1	466	14	B0082995
37	269.8	41.6	426	14	BM856797
38	263.4	40.6	416	14	B0083004
39	263	40.5	479	14	BM741617
40	263	40.5	541	14	BM846805
41	263	40.5	560	14	BM847170
42	263	40.5	594	14	BM763326
43	263	40.5	841	12	BG750836
44	263	40.5	968	14	B0653019
45	263	40.5	993	9	AL513838

ALIGNMENTS

RESULT 1
LOCUS BT65703 728 bp mRNA linear EST 25-SEP-2001
DEFINITION 603046670P1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5186848 5',
ACCESSION BT65703
VERSION BT65703.1 GI:15757281
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 728)
AUTHORS NIH-MGC
TITLE NIH-MGC
JOURNAL http://mgc.nci.nih.gov/
COMMENT Contact: Robert Strausberg, Ph.D.
Unpublished (1999)
National Institutes of Health, Mammalian Gene Collection (MGC)
Email: cgaabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cdna Library Arrayed by: The I.M.A.G.E. Consortium (LMN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMN at:
http://image.llnl.gov
Plate: L1AM11466 row: 1 column: 17
High quality sequence stop: 718.
Location/Qualifiers
1..728

FEATURES
source

db	seq	start	end	score
db	524 AAAAGTCAACCTTGGACACACTGTGAGACC	1	14	583
QY	199 TGCCAGTCCGA	210		
.db	584 AAACAGACTGA	595		

Search completed: April 15, 2003, 15:16:24
Job time : 131.963 secs

THIS MAN IS DEAD

Best Local Similarity 56.1%; Pred. No. 1.3;
Matches 74; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

OY 79 TACACACCACTGCGTCCCTGCAAGGAGGACACTTCCAGATACCTCTCCCGGAGGC 138
DB 464 TGATACCATCTGCGAGCCCTGCGCTTCTTCCAGATGTATCTCTTCCGA 523
OY 139 CCGCTGCGAGCCCGACAGGAGTGTGAGAACCAAGTGTGTGAGAGGACCTCCAGGCAC 198
DB 524 AAAATGTCACTGTGACCAAGCTGTGAGACCAAGACCTGTGTGTGACAGGAGGCAC 583
OY 199 TGCCAGTCCGA 210
DB 584 AACACAGACTGA 595

RESULT 14
AAQ47341
ID AAQ47341 standard; cDNA; 1004 BP.
XX
AC AAQ47341;
XX
DT 07-FEB-1994 (first entry)
XX
DE CD40 coding sequence.
XX
KM Receptor; ligand; B-cell; T-cell; allergy; autoimmunity; antibody;
XX ss.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT CDS 48..876
ET /*lag= a
FT /product= CD40CR Receptor.
XX
PN EP555880-A.
XX
PD 18-AUG-1993.
XX
PE 12-FEB-1993; 93EP-0102279.
XX
PR 14-FEB-1992; 92US-0835799.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
PA (DART-) DARTMOUTH COLLEGE.
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Aruffo AA, Ledbetter JA, Noell R, Stamenkovic I;
PI Noelle R;
XX
XX WPI; 1993-260142/33.
DR P-PSDB; AAR38859.
XX
PT CD40CR receptor and its' ligands - used to inhibit B-cell
XX activation in allergy and auto-immune disease
XX
XX Disclosure: Figure 8a; 21pp; English.
XX
XX The CD40CR receptor is a counter receptor for the CD40 B-cell
XX antigen. It is also a receptor for ligands (sometimes fusion
XX molecules) comprising part of the CD40 protein. A soluble
XX CD40/immunoglobulin fusion protein is able to inhibit helper T-cell
XX mediated B-cell activation by binding to the CD40 receptor on
XX T-cell membranes. Purified receptor provides a means of
XX controlling B-cell activation which may be useful in the treatment
XX of allergy and autoimmune disease.
SQ Sequence 1004 BP; 230 A; 314 C; 259 G; 201 T; 0 other;

Query Match 6.0%; Score 39.2; DB 14; Length 1004;
Best Local Similarity 56.1%; Pred. No. 1.3;
Matches 74; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

OY 79 TACACACCACTGCGTCCCTGCAAGGAGGACACTTCCAGATACCTCTCCCGGAGGC 138
DB 464 TGATACCATCTGCGAGCCCTGCGCTTCTTCCAGATGTATCTCTTCCGA 523
OY 139 CCGCTGCGAGCCCGACAGGAGTGTGAGAACCAAGTGTGTGAGAGGACCTCCAGGCAC 198
DB 524 AAAATGTCACTGTGACCAAGCTGTGAGACCAAGACCTGTGTGTGACAGGAGGCAC 583
OY 199 TGCCAGTCCGA 210
DB 584 AACACAGACTGA 595

RESULT 15
AAT14706
ID AAT14706 standard; cDNA; 1004 BP.
XX
AC AAT14706;
XX
DT 30-OCT-1996 (first entry)
XX
DE Human CD40 antigen cDNA.
XX
KM Cell surface antigen; cloning; immunoselection; immunotherapy;
XX therapy; diagnosis; vector; CD40; COS; lymphocyte; ss.
XX
OS Homo sapiens.
XX
PN US5506126-A.
XX
PD 09-APR-1996.
XX
PE 25-FEB-1988; 88US-0160416.
XX
PR 01-DEC-1992; 92US-0983647.
XX
PR 25-FEB-1988; 88US-0160416.
XX
PR 13-JUL-1989; 89US-0379076.
XX
PR 13-JUL-1990; 90US-0553759.
XX
PR 18-OCT-1993; 93US-0139273.
XX
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Aruffo A, Seed B;
XX
XX WPI; 1996-200279/20.
DR
XX
XX Cloning of cDNA encoding cell surface antigen - useful for isolation
XX of diagnostic and therapeutic proteins
XX
XX Example 8; Fig 16; 79pp; English.
XX
XX A cDNA clone (AAT14714) codes for human cell surface antigen CD40.
XX It was isolated using a rapid immunoselection method in which
XX a cDNA library is constructed in mammalian (e.g. COS) cells using
XX novel expression vectors (see also AAT14702 and AAT14705), and cells
XX expressing the antigen are selected using antibody-coated plates
XX (panning). This immunoselection cloning method, developed to clone
XX genes for cell surface antigens of human lymphocytes (see also
XX CC AAT14703-04 and AAT14706-26), has general appln. Cell surface antigens
XX are obtd. for diagnostic and therapeutic use.
SQ Sequence 1004 BP; 230 A; 297 C; 276 G; 201 T; 0 other;

Query Match 6.0%; Score 39.2; DB 17; Length 1004;
Best Local Similarity 56.1%; Pred. No. 1.3;
Matches 74; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

OY 79 TACACACCACTGCGTCCCTGCAAGGAGGACACTTCCAGATACCTCTCCCGGAGGC 138
DB 464 TGATACCATCTGCGAGCCCTGCGCTTCTTCCAGATGTATCTCTTCCGA 523
OY 139 CCGCTGCGAGCCCGACAGGAGTGTGAGAACCAAGTGTGTGAGAGGACCTCCAGGCAC 198

XX Duda A, Finkel K, Koshy B, Parks KE;
 XX WPI: 2002-154721/20.
 DR P-PSDB; AAM48922.
 XX Novel genetic variants of enolase 3, (beta, muscle) gene useful in
 PT studying expression and function of the protein, and for screening
 PT drugs to treat disorders of glycolytic pathway -
 XX
 PS Claim 20: Fig 1; 90pp; English.
 XX
 CC The present invention provides the protein, cDNA and genomic sequences of
 CC a human enolase 3 (beta, muscle) isoenzyme containing a number of single
 CC nucleotide polymorphisms (SNPs). The sequences can be used to identify
 CC the haplotype of an individual and identify whether particular haplotypes
 CC are linked to certain diseases. The present sequence is the ENO3 gene.
 XX
 SQ Sequence 13308 BP; 2874 A; 3775 C; 3758 G; 2901 T; 0 other;

Query Match 6.1%; Score 39.4; DB 24; Length 13308;
 Best Local Similarity 45.6%; Pred. No. 2.3; Mismatches 166; Indels 0; Gaps 0;
 Matches 139; Conservative 0;

QY 51 CCGAGCTCAAGATGAAGTGGGAAGGTAACAACCACTGCGTCCCTGCAAGCAGAGGC 110
 DB 11006 CCGATGACACTTAAATATACGATCTTACAAAACCTCTGGGCCGAGCAGTGGCTC 11065
 QY 111 ACTTCCAAATACCTCTCTCCCGCCGCGCTGCGACGCCACACCAAGGTGAGAAC 170
 DB 11066 ACTCTTGCATCCAGACACTTTGGAGGCTGAGCGCGGATCAGCTGAGGTAGAGT 11125
 QY 171 AAGGCTGTGTGAGACAGCTCCAGCAGCACTGCCAGTCCGACAACTGCAAAATCCAT 230
 DB 11126 TAGAGACACGCTGCGCAATGTGTGAACCCATCTCTACTAATATGACAAAAATTAC 11185
 QY 231 TAGAGCAGCTGCCCGCAGAGATGTAGAGTGTCTCAAGAGCGTCCGAGGAGAGG 290
 DB 11186 CCGGGTGTGTGACACATGCTGTAGTCCAGCTACTGAGAGGCTAAGGAGAGATC 11245
 QY 291 GACCCATCTGTAGCTGGAAGCTGGAGAGCTCCGAAGGCCATCATCTCTCTACT 350
 DB 11246 GCTTGAACCTGTGAGGAGAGAGGTGAGTCCAGAGTACACCACTGCACTCCAGTCT 11305
 QY 351 TGGTA 355
 DB 11306 GGCCA 11310

RESULT 12
 AAN90617
 ID AAN90617 standard; cDNA; 1004 BP.
 XX
 AC AAN90617;
 XX
 DT 20-DEC-1989 (first entry)
 XX
 DE CD40 cDNA.
 XX
 KW Tyrosine suppressor gene; ASV LTR; human cytomegalovirus AD169 enhancer;
 KW HIV box; immunoselection; immune deficiency diseases; vasculitis;
 KW systemic lupus erythematosus; rheumatoid arthritis; neoplasms;
 XX
 PN EP330191-A.
 XX
 PD 30-AUG-1989.
 XX
 PF 23-FEB-1989; 89EP-0103127.
 XX
 PR 25-FEB-1988; 88US-0160416.
 XX
 PA (GEHO) GEN. HOSPITAL CORP.
 XX

PI Seed B, Allen J, Aruffo A, Camerini D, Laufer L, Oquendo CP,
 PI Simmons D, Stamenkovic I, Stengelin S;
 XX WPI: 1989-250302/35.
 DR
 XX
 XX Rapid immuno:selection cloning - used to clone genes encoding
 PT cell surface antigens associated with mammalian T lymphocytes.
 XX
 PS Disclosure; fig. 17; 69pp; English.
 XX
 CC CD40 encodes a cell surface antigen involved in cell mediated immunity.
 CC This DNA can be expressed in a vector which transforms COS cells. The
 CC vector can isolate any protein, and clones are easy to manipulate.
 XX

SQ Sequence 1004 BP; 230 A; 297 C; 276 G; 201 T; 0 other;

Query Match 6.0%; Score 39.2; DB 10; Length 1004;
 Best Local Similarity 56.1%; Pred. No. 1.3; Mismatches 58; Indels 0; Gaps 0;
 Matches 74; Conservative 0;

QY 79 TAACACCACTGCGCTCCCTGCAAGGCGCACTTCCAGATACCTCTCCCGCCAGCGC 138
 DB 464 TGATACCATCTGCGAGCCCTGCGCAGTGGCTTCTTCCATGTGATCTGCTTTCGA 523
 QY 139 CCGCTGCCAGCCCGCACCAAGGTGTGAGAAACCAAGTCTGTGAGGAGCTCCAGGCAC 198
 DB 524 AAATGTACCTCTGGAACAGCTGTGAGACCAAGACCTGTGTGCAACAGCGAGCAC 583
 QY 199 TGGCAGTCCGA 210
 DB 584 AAMCAGACTGA 595

RESULT 13

AAQ21176
 ID AAQ21176 standard; DNA; 1004 BP.

AC AAQ21176;
 XX

DT 21-MAY-1992 (first entry)
 XX

DE Human CD40 antigen coding sequence.
 XX

KW Cloning technique; cell surface antigen; immunodiagnosis; ss.
 XX

OS Homo sapiens.
 XX

PN WO9201049-A.
 XX

PD 23-JAN-1992.
 XX

PE 15-JUL-1990; 90WO-US04986.
 XX

PR 13-JUL-1990; 90US-0553759.
 XX

PA (GEHO-) GEN HOSPITAL CORP.
 XX

PI Seed B, Aruffo A, Amlot M;
 XX

DR WPI: 1992-056864/07.
 XX

PT New CD53 cell surface antigen and DNA encoding it - for
 PT immuno-therapy and diagnosis of haematopoietic neoplasms, etc.
 XX

PS Example 8; Fig 17; 160pp; English.
 XX

CC Sequences coding for several human antigens were cloned using a new
 CC rapid immunoselection technique (see e.g. AAQ21164). This sequence
 CC encodes the CD40 antigen.
 XX

SQ Sequence 1004 BP; 230 A; 298 C; 275 G; 201 T; 0 other;

Query Match 6.0%; Score 39.2; DB 13; Length 1004;

Db 239 CGAGTGTACACACATGCGAGCTACATTCGTGACTGCCCGCTGGCACTAACGCCGAGCTCAA 298
 Oy 61 AGATGAAGTTGGGAAGGGTAACACACACTGCGTCCCTGGCAAGGAGGCACTT--CCAG 118
 Db 299 ACATGAAGTTGGGAAGGGTAACACACACTGCGTCCCTGGCAAGGAGGCACTTCCAGA 358
 Oy 119 AATACCTCTCCCTCCGAGCGCCGCTGCCAGGCCACACAGGTGTGAGAACCCAGGTCTG 178
 Db 359 AGTACCTCTCCCTCCGAGCGCCGCTGCCAGGCCACACAGGTGTGAGAACCCAGGTCTG 418
 Oy 179 GTGGAGGGACGATCCAGGACCTCCAGTCCAGACACACCTGCAAAAATCCATTAGAGCA 238
 Db 419 GTGGAGGGACGATCCAGGACCTCCAGTCCAGACACACCTGCAAAAATCCATTAGAGCA 478
 Oy 239 CTG 241
 Db 479 CTG 481
 RESULT 9
 ABN41522
 ID ABN41522 standard; DNA; 60 BP.
 AC ABN41522;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human spliced transcript detection oligonucleotide SEQ ID NO:14270.
 XX
 KW Human; mouse; rat; splice transcript; detection; RNA transcript;
 KM splice variant; transcriptome; oligonucleotide library; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200210449-A2.
 XX
 PD 07-FEB-2002.
 XX
 PF 20-JUL-2001; 2001WO-1B01903.
 XX
 PR 28-JUL-2000; 2000US-221607P.
 PR 02-MAY-2001; 2001US-287724P.
 XX
 PA (COMP-) COMPUGEN INC.
 PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
 DR WPI; 2002-257383/30.
 XX
 PT New oligonucleotide libraries comprising oligonucleotides which
 PT selectively hybridize to mRNAs transcribed from a transcription unit of
 PT a genome, useful for detecting tissue-, pathology-, and
 PT developmental-specific genes
 PS
 XX Example 1; SEQ ID 14270; 47bp; English.
 CC The present invention describes oligonucleotide libraries for detecting
 CC messenger RNAs that populate a (sub-)transcriptome, where the
 CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
 CC transcription units that populate a genome. The library comprises
 CC several oligonucleotides, each capable of hybridizing selectively to a
 CC set of messenger RNAs transcribed from a given transcription unit of
 CC the genome, which encodes one or more messenger RNA splice variants.
 CC The oligonucleotide libraries are useful for detecting mRNAs from a
 CC biological sample, in expression profiling studies, in qualitatively or
 CC quantitatively characterizing the corresponding transcriptome, and in
 CC detecting RNA transcripts and splice variants of human or animal
 CC transcriptomes. The libraries may also be used as specialized mini
 CC libraries to detect transcripts of a sub-transcriptome under a
 CC particular biological or pathological state, and so allowing the
 CC detection of tissue- and pathology-specific genes such as those genes
 CC only expressed in specific tissue under a specific pathological
 CC condition; to detect developmental specific genes; and to detect RNA

CC transcripts and splice variants of a transcriptome of a patient suffering
 CC from a particular disorder. ABN27253 to ABN59589 represent
 CC oligonucleotide sequences from rats, humans and mice, which are used in
 CC the exemplification of the present invention.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).
 SO Sequence 60 BP; 18 A; 15 C; 17 G; 10 T; 0 other;
 Query Match 9.28; Score 60; DB 24; Length 60;
 Best Local Similarity 100.0%; Pred. No. 1.9e-06;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 63 ATGAAGTTGGGAAGGGTAACACACACTGCGTCCCTGGCAAGGAGGCACTTCCAGATA 122
 Db 1 ATGAAGTTGGGAAGGGTAACACACACTGCGTCCCTGGCAAGGAGGCACTTCCAGATA 60
 RESULT 10
 AA51857
 ID AA51857 standard; cDNA; 845 BP.
 XX
 AC AA51857;
 XX
 DT 31-OCT-2000 (first entry)
 XX
 DE Porcine CD40 cDNA.
 XX
 KW CD40 antigen; xenograft; CD40-deficient donor; graft rejection;
 KM transgenic pig; reduced immunogenicity; immunosuppressive;
 KW gene therapy; ss.
 XX
 OS Sus scrofa.
 XX
 FH Key Location/Qualifiers
 FT CDS 9..845
 FT /*tag= a
 FT /product= CD40
 XX
 PN WO200039294-A1.
 XX
 PD 06-JUL-2000.
 XX
 PF 22-DEC-1999; 99WO-EP10332.
 XX
 PR 24-DEC-1998; 98GB-0028705.
 PR 10-FEB-1999; 99GB-0002940.
 XX
 PA (NOVS) NOVARTIS AG.
 PA (NOVS) NOVARTIS-ERFINDUNGEN VERN GES MBH.
 PI Bravery C, Rushworth S, Thompson S;
 DR WPI; 2000-452392/39.
 XX
 PT Ameliorating cellular graft rejection in a recipient mammal using
 PT CD40-deficient donor mammal cells, tissues or organs as a xenograft,
 PT comprises inactivation of the CD40 antigen gene with a targeting vector
 PS
 XX Claim 10; Fig 1; 23pp; English.
 CC This porcine cDNA encodes CD40 antigen. A fusion protein consisting of
 CC the extracellular domain of human CD134 coupled to mouse Ig kappa light
 CC chain was able to bind to porcine CD40, indicating that direct signalling
 CC between a porcine graft and the immune cells of a human host is possible
 CC through this molecular pathway. CD40-deficient donor mammal cells,
 CC tissues or organs may be used for xenografting in order to ameliorate
 CC cellular graft rejection in the recipient mammal. In particular, porcine
 CC cells, tissues or organs are genetically modified so that the cell
 CC surface expression of CD40 antigen is reduced. Transgenic pigs may have
 CC an inactivated CD40 antigen gene and a transgene, e.g. human Decay
 CC Accelerating Factor (hDAF) gene. The CD40-deficient donor mammal cells,

CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer.
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.

CC Sequence 2136 BP; 446 A; 706 C; 608 G; 376 T; 0 other;

Query Match 81.8%; Score 531; DB 24; Length 2136;

Best Local Similarity 85.7%; Pred. No. 2,4e-130;

Matches 649; Conservative 0; Mismatches 0; Indels 108; Gaps 1;

1 CGAGTGTACACACTGCGAGCTACTTCTGACTGCGCGCTGCGACTGAGCCGAGCTCAA 60
 579 CGAGTGTACACACTGCGAGCTACTTCTGACTGCGCGCTGCGACTGAGCCGAGCTCAA 638
 61 AGATGAGTTGGAGAGGGTAAACACACACTGCGCTGCTGCGAGCGAGCGACTTCCAGAA 120
 639 AGATGAGTTGGAGAGGGTAAACACACACTGCGCTGCTGCGAGCGAGCGACTTCCAGAA 698
 121 TACTCTCTCCCGCGAGCGCGCTGCGAGCGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 180
 699 TACTCTCTCCCGCGAGCGCGCTGCGAGCGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 758
 181 GGAGGAGCTCCAGGAGCTGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 240
 759 GGAGGAGCTCCAGGAGCTGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 818
 241 GCGCGCGAGAGATGTCA----- 256
 819 GCGCGCGAGAGATGTCAAGGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 878
 257 ----- 256
 879 TCTGCTCTTGGCAGCGCTTCTCTGCTGAGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 938
 257 -----GATGCTGCTCAAGAGCGCTGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 312
 939 ACTGGAGTCTGCTCTCAAGAGCGCTGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 998
 313 CTGGGAGCTCCGAGAGCGCGATCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 372
 999 CTGGGAGCTCCGAGAGCGCGATCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1058
 373 TTTGAGAGATGTTTCCCGAGTATCCACTGAGGCTCCCGAGCGAGCGAGCGAGCGAGCGAG 432
 1059 TTTGAGAGATGTTTCCCGAGTATCCACTGAGGCTCCCGAGCGAGCGAGCGAGCGAGCGAG 1118
 433 GGTGCGCGAGAGAGAGTCTCTGAGACCTGACGAGGAGCGAGCGAGCGAGCGAGCGAGCGAG 492
 1119 GGTGCGCGAGAGAGAGTCTCTGAGACCTGACGAGGAGCGAGCGAGCGAGCGAGCGAGCGAG 1178
 493 GCAGAGCGAGTGGCGCGAGTATCCAAATGAGTATCCGAGCGAGCGAGCGAGCGAGCGAGTAT 552
 1179 GCAGAGCGAGTGGCGCGAGTATCCAAATGAGTATCCGAGCGAGCGAGCGAGCGAGCGAGTAT 1238
 553 CACTGCGAGAGTATCAATGAGAGCGAGTATCCGAGCGAGCGAGCGAGCGAGCGAGCGAGTAT 612
 1239 CACTGCGAGAGTATCAATGAGAGCGAGTATCCGAGCGAGCGAGCGAGCGAGCGAGCGAGTAT 1298
 613 CTTCCAGCTACCCCGAGCTCCATACCCATTTCC 649
 1299 CTTCCAGCTACCCCGAGCTCCATACCCATTTCC 1335

RESULT 5
 ABQ55003
 ID ABQ55003 standard; cDNA; 1594 BP.
 XX
 AC ABQ55003;
 XX
 DT 22-AUG-2002 (first entry)
 XX
 XX Human ovarian antigen HSABJ44 cDNA, SEQ ID NO:883.
 XX
 XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynecological; reproductive; chromosome 12p13;
 KW gene; ss.
 KW
 KW Homo sapiens.
 OS
 OS WO200200677-A1.
 XX
 XX 03-JAN-2002.
 XX
 XX 07-JUN-2001; 2001WO-US18569.
 XX
 XX 07-JUN-2000; 2000US-209467P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Birse CE, Rosen CA;
 XX
 XX WPI, 2002-147878/19.
 XX
 XX P-PSDB; ABP41926.
 DR
 XX
 XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.
 PT ovarian cancer), immune disorders, cardiovascular disorders and
 PT neurological diseases -
 PT
 XX
 PS Claim 1; SEQ ID No 883; 2922bp; English.
 CC
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovarian and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents cDNA encoding a human ovarian antigen of the
 CC invention.


```

QY 1 CGAGTGTACACACTGCGAGCTACTTTCTGACMGGCCGCTGGCACTGAAGCGAGCTCAA 60
DB 579 CGAGGTGTACACACTGCGAGCTACTTTCTGACMGGCCGCTGGCACTGAAGCGAGCTCAA 638
QY 61 AGATGAAGTTGGGAAGGATACACACACTGCGCTCCCTGCAAGGAGGCACTTCCAGAA 120
DB 639 AGATGAAGTTGGGAAGGATACACACACTGCGCTCCCTGCAAGGAGGCACTTCCAGAA 698
QY 121 TACCTCTCTCCCGAGCGCCGCTGCGACGCCACACACAGGTGTGAACCAAGGCTGTGT 180
DB 699 TACCTCTCTCCCGAGCGCCGCTGCGACGCCACACACAGGTGTGAACCAAGGCTGTGT 758
QY 181 GGAGGAGCTCCAGGCACTGCGACAGTCCAGACACAACTTGCACAAAATTCATTAGACCACT 240
DB 759 GGAGGAGCTCCAGGCACTGCGACAGTCCAGACACAACTTGCACAAAATTCATTAGACCACT 818
QY 241 GCGCCGAGAGATGCA----- 256
DB 819 GCGCCGAGAGATGTCAGGAACCAATGCTGATGCTGGCCGCTTCTGCTGCGACACTGCGCTTCT 878
QY 257 ----- 256
DB 879 TCTGCTCTTCCACACCTCTTCTCTCTGCACTCTGGAAGACACACCTTCTCTCTGCGAGAA 938
QY 257 -----GATTCGCTCTCTCAAGAGGCGTCCGAGGAGGAGAACCCATCTCTGCTGAGAG 312
DB 939 ACTGGATTCGCTCTCTCAAGAGGCGTCCGAGGAGGAGAACCCATCTCTGCTGAGAG 998
QY 313 CTGGAGACCTCCGAGGCGCCATCACTTCCCTGACTTGTGACAGCACTGCTACCCAT 372
DB 999 CTGGAGACCTCCGAGGCGCCATCACTTCCCTGACTTGTGACAGCACTGCTACCCAT 1058
QY 373 TTCTGAGATGTTTCCCACTATCCACTGCGCTCCCGAGCCCGAGTTTGGAGCAGG 432
DB 1059 TTCTGAGATGTTTCCCACTATCCCACTGCGCTCCCGAGCCCGAGTTTGGAGCAGG 1118
QY 433 GGTGCGCCGAGAGAGAGTCTCTGAGACCTGACAGGAGCGGAGTTGGAGACCCGAGGA 492
DB 1119 GGTGCGCCGAGAGAGAGTCTCTGAGACCTGACAGGAGCGGAGTTGGAGACCCGAGGA 1178
QY 493 GCAGAGCAGAGTGGCCGACGCTGACCAATGAGATTCATGTACCGCGGAGTCTATGACTAT 552
DB 1179 GCAGAGCAGAGTGGCCGACGCTGACCAATGAGATTCATGTACCGCGGAGTCTATGACTAT 1238
QY 553 CACTGGCAACATCTACATCTACAAATGAGCCAGTACTGGGGGAGACCGGGTCTGTGAGA 612
DB 1239 CACTGGCAACATCTACATCTACAAATGAGCCAGTACTGGGGGAGACCGGGTCTGTGAGA 1298
QY 613 CCTCCAGCTACCCCGGAGCTCCATACCCCATTCCTCC 649
DB 1299 CCTCCAGCTACCCCGGAGCTCCATACCCCATTCCTCC 1335

```

RESULT 3
 ABL64078
 ID ABL64078 standard; DNA; 2136 BP.

AC ABL64078;
 DT 15-MAY-2002 (first entry)
 DE Breast cancer related gene sequence SEQ ID NO:2415.
 XX Human; cancer: colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
 KW gene; ds.
 OS Homo sapiens.
 XX
 PN WO200194629-A2.
 XX
 PD 13-DEC-2001.

XX
 PF 30-MAY-2001; 2001WO-US10838.
 XX
 PR 05-JUN-2000; 2000US-209473P.
 PR 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-233133P.
 PR 18-SEP-2000; 2000US-233617P.
 PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234974P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 26-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 XX
 PA (AVAL-) AVALON PHARM.
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 DR WPI; 2002-188264/24.
 XX
 PT Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set
 XX
 PS Claim 1, SEQ ID 2415; 44pp; English.
 XX
 CC The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,

PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample
PS Claim 1: SEQ ID NO 2135; 298bp; English.
XX
XX
CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pof_sequences.
XX
XX
SQ Sequence 2136 BP; 446 A; 706 C; 608 G; 376 T; 0 other;

Query Match 81.8%; Score 531; DB 24; Length 2136;
Best Local Similarity 85.7%; Pred. No. 2,4e-130;
Matches 649; Conservative 0; Mismatches 0; Indels 108; Gaps 1;

OY 1 CGAGTGACACTGCGAGCTACTTCTGACTCCGCGCTGCGACGTGAACCGAGCTCAA 60
DB 579 CGAGTGACACTGCGAGCTACTTCTGACTCCGCGCTGCGACGTGAACCGAGCTCAA 638
OY 61 AGATGAAGTTGGAGGAGGTAAACAACCACTGCTCCCTGCAAGGAGGAGCACTTCCAGAA 120
DB 639 AGATGAAGTTGGAGGAGGTAAACAACCACTGCTCCCTGCAAGGAGGAGCACTTCCAGAA 638
OY 121 TACCTCTCTCCCGAGCGCGCTGCGACCGCCACACAGGTGAGAACCAAGTCTGGT 180
DB 699 TACCTCTCTCCCGAGCGCGCTGCGACCGCCACACAGGTGAGAACCAAGTCTGGT 758
OY 181 GGAGGAGCTCCAGGCACTGCGAGTCCGACACAACTGCAAAAATTCATTAGAGCACT 240
DB 759 GGAGGAGCTCCAGGCACTGCGAGTCCGACACAACTGCAAAAATTCATTAGAGCACT 818
OY 241 GCCCCAGAGATGTCA----- 256
DB 819 GCCCCAGAGATGTCAAGAACATGCTGATGCGGCCGTTCTGCTGCACTGGGCTTCT 878
OY 257 ----- 256
DB 879 TCTGCTCTTTCGCAACCGTCTTCTGCTGATCTGGAAGAGCCACCTTCTCTGCAAGAA 938
OY 257 -----GATCGCTGCTCAAGAGGCGTCCGAGGAGAGAGGAGCCCAATCTGAGTGAAG 312
DB 939 ACTGCGATCTGCTCAAGAGGCGTCCGAGGAGAGAGGAGCCCAATCTGAGTGAAG 998
OY 313 CTGGAGACCTCCGAGGCGCCATCCATCTTCCCTGACTTGATACAGCCACTGTACCCAT 372
DB 999 CTGGAGACCTCCGAGGCGCCATCCATCTTCCCTGACTTGATACAGCCACTGTACCCAT 1058
OY 373 TTCTGGAGATGTTTCCCAATATTCATCTGGGCTCCCGAGGCCCACTTTTGAAGGAGG 432
DB 1059 TTCTGGAGATGTTTCCCAATATTCATCTGGGCTCCCGAGGCCCACTTTTGAAGGAGG 1118
OY 433 GGTGCGCCACAGCAGAGTCTCTGACCTGACAGGAGGAGCCGAGTGAACCCGGGGA 492
DB 1119 GGTGCGCCACAGCAGAGTCTCTGACCTGACAGGAGGAGCCGAGTGAACCCGGGGA 1178
OY 493 GCAGAGCCAGAGTGGCCACGAGTGAACCAATGAGTATGATGATGAT 552
DB 1179 GCAGAGCCAGAGTGGCCACGAGTGAACCAATGAGTATGATGATGAT 1238
OY 553 CACGCGCAATATCTACATTCACAATGAGAGTACTGGGGAGACCAAGGGGTCTGTGAGA 612
DB 1238 CACGCGCAATATCTACATTCACAATGAGAGTACTGGGGAGACCAAGGGGTCTGTGAGA 612

DB 1239 CACTGGCAACATCTACATCTACATGAGACCACTGCGGGGAGCACCGGGTCTTGAGA 1298
OY 613 CCTCCAGCTACCCCGCAACCTCCATCCATCCATCC 649
DB 1299 CCTCCAGCTACCCCGCAACCTCCATCCATCCATCC 1335

RESULT 2
ID ABK64502
ID ABK64502 standard; DNA; 2136 BP.
AC ABK64502;
XX
XX 18-JUN-2002 (first entry)
DT
XX
DE Human benign prostatic hyperplasia gene #397.
XX
XX Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.
XX
OS Homo sapiens.
XX
XX WO200212440-A2.
XX
PD 14-FEB-2002.
XX
PF 07-AUG-2001; 2001WO-US24708.
XX
PR 07-AUG-2000; 2000US-22323P.
PR 05-JUN-2001; 2001US-0873319.
XX
PA (GENE-) GENE LOGIC INC.
PA (NISB) JAPAN TOBACCO INC.
XX
PI Munger WE, Kulikarni P, Getzenberg RH, Waga I, Yamamoto J;
XX
DR WPI; 2002-257476/30.
XX
XX
PT Identifying drugs for and diagnosing benign prostatic hyperplasia, by
PT detecting expression levels of one or more genes in prostate cells from
PT patient that are differentially regulated compared to normal prostate
PT cells -
XX
XX
PS Disclosure; Page 229-230; 444pp; English.
XX
XX The invention relates to a method of diagnosing (I) the onset or
XX progression of benign prostatic hyperplasia (BPH), or screening (II) for
XX or identifying an agent that modulates the onset or progression of BPH.
XX The method is based on changes in gene expression in BPH tissue isolated
XX from patients exhibiting different clinical states of prostate
XX hyperplasia as compared to normal prostate tissue. (I) comprises
XX detecting the expression levels of one or more genes in prostate cells
XX from the subject that are differentially regulated compared to normal
XX prostate cells. (II) comprises preparing a first gene expression profile
XX of BPH cells or BPH-like cell population, exposing the cells to the
XX agent, preparing a second gene expression profile of the agent exposed
XX cells, and comparing the first and second gene expression profiles.
XX (I) is useful for diagnosing the onset or progression of BPH. (II) is
XX useful for identifying an agent that modulates the onset or progression
XX of BPH. The methods are useful to present information identifying
XX the expression level in a tissue or cells, by comparing the expression
XX level of genes given in the specification in the tissue or cells to the
XX level of expression of gene in the database, and displaying the
XX expression levels of at least one gene in the tissue or cell sample
XX compared to the expression level in BPH. Agents using (II) are useful for
XX treating BPH or prostate cancer. ABK64106-ABK64860 represent human
XX benign prostatic hyperplasia gene sequences of the invention.
XX
SQ Sequence 2136 BP; 446 A; 706 C; 608 G; 376 T; 0 other;

Query Match 81.8%; Score 531; DB 24; Length 2136;
Best Local Similarity 85.7%; Pred. No. 2,4e-130;
Matches 649; Conservative 0; Mismatches 0; Indels 108; Gaps 1;

Run on: April 15, 2003, 13:05:15 ; Search time 112.963 seconds
(without alignments)

Title:	US-09-917-372-2_COPY_629_1277
Perfect score:	649
Sequence:	1 cgaagtacacactgcgagc.....aacctccatcacccattccc 649

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

```
Searched: 2185239 seqs, 1125999159 residues
Total number of hits satisfying chosen parameters: 4370478
```

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_genseq_11002

1:	/SIDS2/gcgdata/genseq/genseqn-emb1/NA1981.DAT.*
2:	/SIDS2/gcgdata/genseq/genseqn-emb1/NA1981.DAT.*
3:	/SIDS2/gcgdata/genseq/genseqn-emb1/NA1982.DAT.*
4:	/SIDS2/gcgdata/genseq/genseqn-emb1/NA1983.DAT.*
5:	/SIDS2/gcgdata/genseq/genseqn-emb1/NA1984.DAT.*
6:	/SIDS2/gcgdata/genseq/genseqn-emb1/NA1985.DAT.*
7:	/SIDS2/gcgdata/genseq/genseqn-emb1/NA1986.DAT.*
8:	/SIDS2/gcgdata/genseq/genseqn-emb1/NA1987.DAT.*
9:	/SIDS2/gcgdata/genseq/genseqn-emb1/NA1988.DAT.*
10:	/SIDS2/gcgdata/genseq/genseqn-emb1/NA1989.DAT.*
11:	/SIDS2/gcgdata/genseq/genseqn-emb1/NA1990.DAT.*
12:	/SIDS2/gcgdata/genseq/genseqn-emb1/NA1991.DAT.*
13:	/SIDS2/gcgdata/genseq/genseqn-emb1/NA1992.DAT.*
14:	/SIDS2/gcgdata/genseq/genseqn-emb1/NA1993.DAT.*
15:	/SIDS2/gcgdata/genseq/genseqn-emb1/NA1994.DAT.*
16:	/SIDS2/gcgdata/genseq/genseqn-emb1/NA1995.DAT.*
17:	/SIDS2/gcgdata/genseq/genseqn-emb1/NA1996.DAT.*
18:	/SIDS2/gcgdata/genseq/genseqn-emb1/NA1997.DAT.*
19:	/SIDS2/gcgdata/genseq/genseqn-emb1/NA1998.DAT.*
20:	/SIDS2/gcgdata/genseq/genseqn-emb1/NA1999.DAT.*
21:	/SIDS2/gcgdata/genseq/genseqn-emb1/NA2000.DAT.*
22:	/SIDS2/gcgdata/genseq/genseqn-emb1/NA2001A.DAT.*
23:	/SIDS2/gcgdata/genseq/genseqn-emb1/NA2001B.DAT.*
24:	/SIDS2/gcgdata/genseq/genseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query	Match	Length	DB	ID	Description
No.							
1	531	81.8	2136	24	ABN95637	Gene #2135 used to	
2	531	81.8	2136	24	ABK64502	Human benign prost	
3	531	81.8	2136	24	ABJ64078	Breast cancer rela	
4	531	81.8	2136	24	ABJ64384	Stomach cancer rel	
5	470	72.4	1594	24	ABO50503	Human ovarian antil	
6	253.8	39.1	281	24	ABJ82623	Human ovarian canc	
c 7	251.2	38.7	289	24	ABJ82608	Human ovarian canc	
8	208.2	32.1	481	24	ABJ81903	Human ovarian canc	
9	60	9.2	60	24	ABN41522	Human spliced trian	

10	40.8	6.3	845	21	AA551857	Porcine CD40 CDNA.
11	39.4	6.1	13308	24	AAK98531	Human CD40 CDNA. AAN90
12	39.2	6.0	1004	19	AAK90617	Human CD40 antigen
13	39.2	6.0	1004	13	AAO21176	CD40 coding sequen
14	39.2	6.0	1004	14	AAO47341	Human CD40 antigen
15	39.2	6.0	1004	17	AAAT14706	Human CD40 antigen
16	39.2	6.0	1004	19	AAV63454	CD40 CDNA. Uniden
17	39.2	6.0	1004	20	AAZ20296	Human CD40 CDNA. H
18	39.2	6.0	1004	20	AAZ23432	Human CD40 antigen
19	39.2	6.0	1004	20	AAZ40936	Human CD40 antigen
20	39.2	6.0	1004	20	AAV81198	Human cell surface
21	39.2	6.0	1004	21	AAV50590	Human CD40 nucleot
22	39.2	6.0	1004	21	AAZ47784	Human lymphocyte c
23	39.2	6.0	1004	22	AA503189	CD40 B-cell antigen
24	39.2	6.0	1004	22	AAAL7328	Human immune/haema
25	39	6.0	1244	22	AAK74605	Human immune/haema
26	39	6.0	15296	22	AAK74605	Human cardiovascular
27	38.8	6.0	15044	22	AA536290	Human cardiovascular
28	38.8	5.9	15046	22	AA536291	Human cardiovascular
29	38.6	5.9	765	21	AAZ55529	Human cardiovascular
30	38.6	5.9	765	21	AAZ55530	Canine mature CD40
31	38.6	5.9	822	21	AAZ55528	Canine mature CD40
32	38.6	5.9	822	21	AAZ55528	Canine CD40 CDNA c
33	38.6	5.9	1425	21	AAZ55525	Canine CD40 CDNA c
34	38.6	5.9	1425	21	AAZ55525	Canine CD40 CDNA.
35	37.4	5.8	6536	19	AAV04471	Canine CD40 CDNA c
36	37.4	5.8	6536	19	AAV04471	Mudarellia beta-1i
37	36.6	5.6	333	20	AAV72224	Nudarellia beta-1i
38	36.6	5.6	459	20	AAK84657	Human anti-Grp1b/I
39	36.4	5.6	459	21	AAAC6757	Human TNFR superia
40	36.4	5.6	767	20	AAK23419	Human TNFR soluble
41	36.4	5.6	813	20	AAZ23377	Human APO6 DNA. H
42	36.4	5.6	813	21	AAV75999	Human mELINT #1 nu
43	36.4	5.6	813	21	AAAB8730	DNA encoding a mat
44	36.4	5.6	813	21	AAAB8730	Human FAS ligand i
45	36.4	5.6	813	22	AAAS1077	Mature human FLINT
					AAAD07380	Human mature fas l

RESULT	1
ID	ABN95637
ABN95637	standard; DNA; 2136 BP.
AC	ABN95637;
XA	13-AUG-2002 (first entry)
DT	
DE	Gene #2135 used to diagnose liver cancer.
XX	Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
XX	metastatic liver tumor; cytostatic; expression profile; disease state;
KW	disease progression; drug toxicity; drug efficacy; drug metabolism.
KW	
XX	
OS	Homo sapiens.
XX	
PN	WO200229103-A2.
XX	
PD	11-APR-2002.
XX	
PF	02-OCT-2001; 2001MO-US30589.
XX	
PR	02-OCT-2000; 2000US-237054P.
XX	
PA	(GENE-) GENE LOGIC INC.
XX	
PI	Horne D, Alvares C, Peres-da-Silva S, Vockley JG;
XX	
DR	WPI; 2002-426119/45.
XX	
PT	Diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma or metastatic liver tumor in a patient,

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

Matches 474; Conservative 0; Mismatches 53; Indels 7; Gaps 5;

```
QY 76 CCCGGCCCTGGGGTGCACATCGGCGCTAGTCCCGTCCAGGCTCTGGGCTCGGGCAGCC 135
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1  CCGGGCTTGGGGTGCACATCGGCGCTAGTCCCGTCCAGGCTCTGGGCTCGGGCAGCC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 136 GCGGCGCACCGCTGCGGAGAGCTGCGGCTCTCCCTCCAGGCCCCCAGGTGCT 195
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61  G-GGCCAMCGCTGCCAGAGAGCTGCGGCTCTCTTTTTCCTCGGCTCTTTTTC 119
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 196 GCGGCGCTGGCGGAGTGGCGGCATG--CTCGTGCCTTGGGCGACCTCTGCCCGGCT 253
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 120 CCGCTCTGGCGAGCTTGGCGGGCTCATGGCTCTGCTCTTGGGGGACCTCTGCCGGGCT 179
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 254 GGCCTGGGGGCTCTGTGCTGGGCTCTTCGCGCTCTCGGAGCATCGACGCCAGGC 313
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 180 GGCCTGGGGCG-CTCTGGTGTGGGCTCTTCGCGGCTCTCGGAGCATCGACGCCAG-- 236
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 314 GGTGCTCCATATGCGCGGAGAACAGACCTGCAGGGAGCCAGGAAAGGAATACTATGA 373
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 237 CGGTGCTCATATGCGTGGAGAACAGACCTGCAGGGAGCCAGGAAAGGAATACTATGA 296
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 374 GCGCCAGCACCGCATCTGTGCTGCCGCTGCCGCGCAGGCACTATGTCTAGCTAAATG 433
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 297 GCGCCAGCACCGCATCTGTGCTGCCGCTGCCGCGCA-GCAGCTATGTCTAGCTAAATG 355
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 434 TAGCCGCAATCCGGAGACACAGTTTGTGCCACATGTGCGAGAAATTCCTACAAAGACACTG 493
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 356 TAGCCGCAATCCGGAGACACAGTTTGTGCCACATGTGCGAGAAATTCCTACAAAGACACTG 415
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 494 GAACTACTGACATCTGCCAGCTGTGCCGCCCTGTGACCCAGTGGGCTCGAGGA 553
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 416 GAACTACTGACATCTGCCAGCTGTGTGCCGCCCTGTGACCCAGTGGGCTCGAGGA 475
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 554 GATTGCCCGCTGCACAAGCAAAACGGAAGACCCAGTGGCGCTGCCAGCGGGGAT 607
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 476 GATTGCCCGCTGCACAAGCAAAACGGAAGACCCAGTGGCGCTGCCAGCGGGGAT 529
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: April 15, 2003, 18:45:36
Job time : 150.101 secs

QY 287 GCTCTGGGACGATCGCAGCCCGGCGGTGCTCCATATGCTGAGAGAACGACCTG 346
 |||||
 Db 180 GCTCTGGGACGATCGCAGCCCGGCGGTGCTCCATATGCTGAGAGAACGACCTG 239
 QY 347 CAGGACACGAGAAAGAAATACTATAGCCCGGACACCGCATCTGCTGCCCGGCGCC 406
 |||||
 Db 240 CAGGACACGAGAAAGAAATACTATAGCCCGGACACCGCATCTGCTGCCCGGCGCC 299
 QY 407 GCAGGACACCTATGCTCAGCTAAATGTAGCCCATCCGGACACAGTTTGTGCCCATG 466
 |||||
 Db 300 GCAGGACACCTATGCTCAGCTAAATGTAGCCCATCCGGACACAGTTTGTGCCCATG 359
 QY 467 TGCCGAGAAATTCCTACACGACGACTGGAACCTACCTGACCATCTGCCAGCTTGCCGCC 526
 |||||
 Db 360 TGCCGAGAAATTCCTACACGACGACTGGAACCTACCTGACCATCTGCCAGCTTGCCGCC 419
 QY 527 CTGTGACCCGATGATGGCGCTCGAGAGATGCGCCCTGACACGAAAGGAGACCCA 586
 |||||
 Db 420 CTGTGACCCGATGATGGCGCTCGAGAGATGCGCCCTGACACGAAAGGAGACCCA 479
 QY 587 GTGCCGCTGCCAGCCGGGATGTTCTGTGCTGCTGGGCGCTCGAGTG 634
 |||||
 Db 480 GTGCCGCTGCCAGCCGGGATGTTCTGTGCTGCTGGGCGCTCGAGTG 527

RESULT 11

US-09-867-701-4881
 ; Sequence 4881, Application US/09867701
 ; Patent No. US20020132237A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Agiate, Paul A.
 ; APPLICANT: Jones, Robert
 ; APPLICANT: Harlocker, Susan L.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
 ; FILE REFERENCE: 210121.497
 ; CURRENT APPLICATION NUMBER: US/09/867,701
 ; CURRENT FILING DATE: 2001-05-29
 ; NUMBER OF SEQ ID NOS: 10912
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4881
 ; LENGTH: 481
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 US-09-867-701-4881

Query Match 21.9%; Score 433.6; DB 10; Length 481;
 Best Local Similarity 96.5%; Pred. No. 5.8e-114;
 Matches 465; Conservative 0; Mismatches 14; Indels 3; Gaps 2;

QY 390 TGCTGCTCCCGCTGCCCGGACGACGACTATGTCTCAGCTAAATGTACCCGATCCGGGAC 449
 |||||
 Db 1 TGCTGCTCCCGCTGCCCGGACGACGACTATGTCTCAGCTAAATGTACCCGATCCGGGAC 59
 QY 450 ACAATTTGTGACATGTGGCGAGAAATTCACACGACGACTGGAATCTACCTGACCATC 509
 |||||
 Db 60 ACAATTTGTGACATGTGGCGAGAAATTCACACGACGACTGGAATCTACCTGACCATC 119
 QY 510 TGCCAGCTGTGCGCCCGCTGTGACCCAGTGTATGGGCTCGAGAGATTGCCCCCTGACACA 569
 |||||
 Db 120 TGCCAGCTGTGCGCCCGCTGTGACCCAGTGTATGGGCTCGAGAGATTGCCCCCTGACACA 179
 QY 570 AGCAAAAGGAGAACCCAGTGGCGCTGCCAGCGGGAAATGTTCTGTGCTGCTGGGCGCTC 629
 |||||
 Db 180 AGCAAAAGGAGAACCCAGTGGCGCTGCCAGCGGGAAATGTTCTGTGCTGCTGGGCGCTC 239
 QY 630 GAGTGTACACACTGCGAGCTACTTCTGACGCGCCGCTGACCTGGAAGCGAGCTCAAA 689
 |||||
 Db 240 GAGTGTACACACTGCGAGCTACTTCTGACGCGCCGCTGACCTGGAAGCGAGCTCAAA 299
 QY 690 GATGAAGTTGGGAGGTTACACACACTGCTGCCCTGCAAGGAGGCGACTT--CCAGA 747
 |||||

Db 300 GATGAAGTTGGGAGGTTACACACCGCTGCTGCACTGCGAAGGAGGCGACTTTCAGAA 359
 QY 748 ATAACCTCTCTCCCGACAGCGCCGCTGCGACGCCACACAGAGTGTGAAACCAAGTCTGG 807
 |||||
 Db 360 GTAACCTCTCTCCCGACAGCGCCGCTGCGACGCCACACAGAGTGTGAAACCAAGTCTGG 419
 QY 808 TGGAGGACACTCAGGACACTGGCCAGTGGGACCAACACCTGCAAAAATCATTTAGAGCCAC 867
 |||||
 Db 420 TGGAGGACACTCAGGACACTGGCCAGTGGGACCAACACCTGCAAAAATCATTTAGAGCCAC 479
 QY 868 TG 869
 ||
 Db 480 TG 481

RESULT 12

US-09-907-372-18/c
 ; Sequence 18, Application US/09907372
 ; Patent No. US20020068242A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lal, Preeti G.
 ; APPLICANT: Warren, Bridget A.
 ; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
 ; FILE REFERENCE: PC-0050 US
 ; CURRENT APPLICATION NUMBER: US/09/907,372
 ; CURRENT FILING DATE: 2001-07-27
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 18
 ; LENGTH: 536
 ; TYPE: DNA
 ; ORGANISM: Macaca fascicularis
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; OTHER INFORMATION: Incyte ID No. US20020068242A1 703678967J1
 US-09-907-372-18

Query Match 21.7%; Score 431; DB 10; Length 536;
 Best Local Similarity 94.6%; Pred. No. 3.3e-113;
 Matches 457; Conservative 0; Mismatches 25; Indels 1; Gaps 1;

QY 1500 GGGGCGCTGAGTAGGGCCCGGGGAGACAGAGCCCTTAAGGATTAGGCTCAGACACTCTG 1559
 |||||
 Db 531 GGGGCGCTGAGTAGGGCCCGGGGAGACAGAGCCCTTAAGGAGGAGACAGACCCCGG 472
 QY 1560 AGAGCAGGTGGGACGTGGGTACGGTCCCTCCACAGACTCCCTACGCTGAG 1619
 |||||
 Db 471 AGAGCAGGTGGGACGTGGGTACGGTCCCTCCACAGACTCCCTACGCTGAG 412
 QY 1620 CAACCTGAGGCTCCCGGAGACACCCACCCCTGCGGGGCTGCTCAGGCTCAGGACAG 1679
 |||||
 Db 411 CAACCTGAGGCTCCCGGAGACACCCACCCACCTGCGGGGCTGCTCAGGCTCAGGACAG 352
 QY 1680 GACAGGCGACATATACCAACTGCTGCCACTACAGACGCGCCGACCGGAGCAGC 1739
 |||||
 Db 351 GACAGGCGACATATACCAACTGCTGCCACTATGAGCACCGCCACCGGAGCAGCGCACT 292
 QY 1740 GAGGAGACCCGACACGCTGACCTGCAAGGACGTCACGCGGCCCTCTAAAGATGCTGG 1799
 |||||
 Db 291 GAGGAGACCCGACACGCTGACCTGCAAGGACGTCACGCGGCCCTCTAAAGATGCTGG 232
 QY 1800 TGCTCATCCCAAGCTTCAGAGACCCCTTGGGGTTCACACTTCACGCTGAGCTGAGTAG 1859
 |||||
 Db 231 TGCTCATCCCAAGCTTCAGAGACCCCTTGGGGTTCACACTTCACGCTGAGTAG 172
 QY 1860 ACCCTGATAGATGAATTAATTAATGAGGAGACGCTCTCTTCCCTCCCTCTAAGAGAG 1919
 |||||
 Db 171 ACCCTGATAGATGAATTAATTAATGAGGAGACGCTCTCTTCCCTCCCTCTAAGAGAG 112
 QY 1920 GAAGGAGTCATTAACAACTAGGGGTTGGGTAGATTCCTAGGTATGGGAGAGATT 1979
 |||||
 Db 111 GAAGGAGTCATTAACAACTA--GGGGTGGGTAGATTCCTAGGTATGGGAGAGATT 53

-09-907-372-8

Best Local Similarity	100.0%	Pred. No. 2,8e-140;							
Matches	526;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;

QY 33 CTGAGCTCCGGCATGGAGCCCTGAGAGGCCCGGCTGGCCCTCCCGGCCCTGGGGTGA 92
|||||

Db 1441 AGGGGCCCCAGGACCAATTATATACCCATGATGACGAGTCTGTAGAAAAAGCAGAA 1500
Oy 1443 AGGGGGGACAAAGGACCTCTTCCCTTGAAGCTGCTCCCTGCCCAGATGGATTCACAGG 1502
Db 1501 AGGGGGGACAAAGGACCTCTTCCCTTGAAGCTGCTCCCTGCCCAGATGGATTCACAGG 1560
Oy 1503 GCGTGAATAGGCCCCGGGGAAAGCAGAGCCCTAAGGGATTAAAGGCTCAGACACCTCTGAA 1562
Db 1561 GCGTGAATAGGCCCCGGGGAAAGCAGAGCCCTAAGGGATTAAAGGCTCAGACACCTCTGAA 1620
Oy 1563 GCGTGAATAGGCCCCGGGGAAAGCAGAGCCCTAAGGGATTAAAGGCTCAGACACCTCTGAA 1622
Db 1621 GCGTGAATAGGCCCCGGGGAAAGCAGAGCCCTAAGGGATTAAAGGCTCAGACACCTCTGAA 1680
Oy 1623 ACCTGAGGACCTCCCGGACAGCCACCCCTGGGGCTGCTAGCTCAGAGCAGGAC 1682
Db 1681 ACCTGAGGACCTCCCGGACAGCCACCCCTGGGGCTGCTAGCTCAGAGCAGGAC 1740
Oy 1683 AGGGACATGATACCACTGCTGCCCCACTACAGCAGCCGCGACGGAGCAGGACCGAG 1742
Db 1741 AGGGACATGATACCACTGCTGCCCCACTACAGCAGCCGCGACGGAGCAGGACCGAG 1800
Oy 1743 GGAGCCGACACAGCTGCTACAGAGAGCTACGAGGAGGAGGAGGAGGAGGAGGAGGAG 1802
Db 1801 GGAGCCGACACAGCTGCTACAGAGAGCTACGAGGAGGAGGAGGAGGAGGAGGAGGAG 1860
Oy 1803 TCATCCCAAGCTTACAGAGACCCCTTGGGGTTCACACTTCACTGAGTGAAGTAGAGC 1862
Db 1861 TCATCCCAAGCTTACAGAGACCCCTTGGGGTTCACACTTCACTGAGTGAAGTAGAGC 1920
Oy 1863 CTGCAATGAATGAATATATAGGAGAGAGCTCTTCCCTCCCTCTCTAGAGAGAGAA 1922
Db 1921 CTGCAATGAATGAATATATAGGAGAGAGCTCTTCCCTCCCTCTCTAGAGAGAGAA 1980
Oy 1923 AGGAGTCAATTAACAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1982
Db 1981 AGGAGTCAATTAACAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2040

RESULT 5
US-09-907-372-12/c
: Sequence 12, Application US/09907372
: Patent No. US20020068242A1
: GENERAL INFORMATION:
: APPLICANT: Lal, Preeti G.
: APPLICANT: Warren, Bridget A.
: TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
: FILE REFERENCE: PC-0050 US
: CURRENT APPLICATION NUMBER: US/09/907,372
: CURRENT FILING DATE: 2001-07-27
: NUMBER OF SEQ ID NOS: 20
: SOFTWARE: PERL Program
: SEQ ID NO 12
: LENGTH: 862
: TYPE: DNA
: ORGANISM: Homo sapiens
: NAME/KEY: misc-feature
: OTHER INFORMATION: Incyte ID No. US20020068242A1 857691871
US-09-907-372-12

Query Match 32.5%; Score 645; DB 10; Length 862;
Best Local Similarity 92.4%; Pred. No. 3.8e-174;
Matches 790; Conservative 0; Mismatches 50; Indels 15; Gaps 10;

Oy 1131 GTGGGCCACGGTACCAATGCGATTCATGTCACCGGGGCTATGACTATCACTGGCAC 1190
Db 833 GTGGGCCCGGGTACCAAGGACATAGGT-ACCGGGGGTTCAAGAACATTCACGGGCAAA 795
Oy 1191 A--TCTACATCTACATGAGACA--GTACTGGGGGAGCACCGGGTCTGTGAGACTCC 1246
Db 794 ACTGCAATTAACAAGGACCAAGTACTGGGGAGAGCACCGGGTCTGTGAGACTCC 735

Oy 1247 AGTACCCCGGACACC--TCATATACCCATTCCTCCGAAAGGGGAGACCTTGCCCTCCGGG 1304
Db 724 AGGCTACCCCGGACACCCTTTCATATACCCATTCCTCCGAAAGGGGAGACCTTGCCCTCCGGG 675
Oy 1305 CTCTCTACACCCGACAGAGAGATGG--CAAGGCTTGGCACCCTAGCGAGACAGAC--AC 1361
Db 674 CTCTCTACACCCGACAGAGAGATGGGCAAGGCTTGGCACCCTAGCGAGAGAGAGAAC 615
Oy 1362 TGTGGTGGCACACCTCTTACAGAGGGGCCCCAAGAACCAATTATACCATATGATGATG 1421
Db 614 TGTGGTGGCACACCTCTTACAGAGGGGCCCCAAGAACCAATTATACCATATGATGATG 555
Oy 1422 AGTCTGAAAAAGGACGAGAAGAGGGGGGACAGAGGACCTTCTCTTGAAGCTGCCCT 1481
Db 554 AGTCTGAAAAAGGACGAGAAGAGGGGGGACAGAGGACCTTCTCTTGAAGCTGCCCT 495
Oy 1482 GCGCCAGTGGGATTCACAGGGGGGCTGATAGGGGCCCCGGGGAAGCAGACCTTAAGGGATT 1541
Db 494 GCGCCAGTGGGATTCACAGGGGGGCTGATAGGGGCCCCGGGGAAGCAGACCTTAAGGGATT 435
Oy 1542 AAGGCTCAGACACCTCTGAGAGACAGTGGGACCTGCTGGGTACGGTCCCTCCACAGA 1601
Db 434 AAGGCTCAGACACCTCTGAGAGACAGTGGGACCTGCTGGGTACGGTCCCTCCACAGA 375
Oy 1602 CTCTCTCTACTGCTGACCAAACTGAGGCTCCCGGACAGACCCACCCCTGCGGGG 1661
Db 374 CTCTCTCTACTGCTGACCAAACTGAGGCTCCCGGACAGACCCACCCCTGCGGGG 316
Oy 1662 TGCTAGGCTCAGAGGACAGGAGGACATGATATCAACTGCTGCTCCACTACAGCAGCC 1721
Db 315 TGCTAGGCTCAGAGGAGGAGGAGGACATGATATCAACTGCTGCTCCACTACAGCAGCC 256
Oy 1722 GCACCGGACACGGG--ACCGAGGAGGAGCCGACACAGGTCACCTGCAAGAGAGTCAAGGG 1780
Db 255 GCACCGGAGACAGGCAACGAGGAGGAGCCGACAGGTCACCTGCAAGAGAGTCAAGGG 196
Oy 1781 CCCTTAAGGATTTGTGTGTCTCATCCCAAGCTTCTGAGAGACCTTTGGGGTTCCAC 1840
Db 195 CCCT-TAAGGATTTGTGTGTCTCATCCCAAGCTTCTGAGAGACCTTTGGGGTTCCAC 137
Oy 1841 TTCAGTGTGAGTGTGAGTACCTGATGATGAATTAATATAGGAGAGAGCTCTCTCC 1900
Db 136 TTCAGTGTGAGTGTGAGTACCTGATGATGAATTAATATAGGAGAGAGCTCTCTCC 77
Oy 1901 CTCCTCTCTAGAGAGAGGAAAGGAGTCAATTAACA--CTAGGGGCTTGGGTAGATT 1958
Db 76 CTCCTCTCTAGAGAGAGGAAAGGAGTCAATTAACAAGTGTGGGTAGATT 17
Oy 1959 CCTAGTATGGGAA 1973
Db 16 CCTAGTATAGGGGA 2

RESULT 6
US-09-907-372-7/c
: Sequence 7, Application US/09907372
: Patent No. US20020068242A1
: GENERAL INFORMATION:
: APPLICANT: Lal, Preeti G.
: APPLICANT: Warren, Bridget A.
: TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
: FILE REFERENCE: PC-0050 US
: CURRENT APPLICATION NUMBER: US/09/907,372
: CURRENT FILING DATE: 2001-07-27
: NUMBER OF SEQ ID NOS: 20
: SOFTWARE: PERL Program
: SEQ ID NO 7
: LENGTH: 651
: TYPE: DNA
: ORGANISM: Homo sapiens
: NAME/KEY: misc-feature
: OTHER INFORMATION: Incyte ID No. US20020068242A1 7716364H1

QY	1743	GGAGCGCCGACACGGGTACACCTGCAGAGAGACGTACGGGGCCCTCTAAAGATTCTGTGGTGC	1802
Db	1801	GGAGCGCCGACACGGGTACACCTGCAGAGAGACGTACGGGGCCCTCTAAAGATTCTGTGGTGC	1860
QY	1803	TCATCCCGACCTTCAGAGACCTTTGGGGGTTCCACACTTCACGTGACGTGAGTACAGACC	1862
Db	1861	TCATCCCGACCTTCAGAGACCTTTGGGGGTTCCACACTTCACGTGAGTACGTGAGTACAGACC	1920
QY	1863	CTGCATGAGATGAAATTATAGAGAGAGACGCTCTCTCCCTCCCTCTAGAGAGAGAA	1922
Db	1921	CTGCATGAGATGAAATTATAGAGAGAGACGCTCTCTCCCTCCCTCTAGAGAGAGAA	1980
QY	1923	AGGGAGTCAATTAACACATGAGGGGCTTGGGTAGATTCTCAGGTATGGGAGAGTTTGG	1982
Db	1981	AGGGAGTCAATTAACACATGAGGGGCTTGGGTAGATTCTCAGGTATGGGAGAGTTTGG	2040
RESULT 4			
US-09-880-107-2135			
Sequence 2135, Application US/09880107			
Patent No. US20020142981A1			
GENERAL INFORMATION:			
APPLICANT: Horne, Darci T.			
APPLICANT: Vockley, Joseph G.			
APPLICANT: Scherf, Uwe			
APPLICANT: Gene Logic, Inc.			
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer			
FILE REFERENCE: 44921-5028-WO			
CURRENT APPLICATION NUMBER: US/09/880.107			
CURRENT FILING DATE: 2001-06-14			
PRIORITY FILING DATE: US 60/211,379			
PRIORITY FILING DATE: 2000-06-14			
PRIORITY APPLICATION NUMBER: US 60/237,054			
PRIORITY FILING DATE: 2000-10-02			
NUMBER OF SEQ ID NOS: 3950			
SOFTWARE: PatentIn Ver. 2.1			
SEQ ID NO 2135			
LENGTH: 2136			
TYPE: DNA			
ORGANISM: Homo sapiens			
FEATURE:			
OTHER INFORMATION: Genbank Accession No. US20020142981A1 L04270			
US-09-880-107-2135			
Query Match			
Best Local Similarity 91.3%; Score 1809.2; DB 10; Length 2136;			
Matches 1929; Conservative 0; Mismatches 3; Indels 108; Gaps			
QY	51	GCCCTGAGAGCCCGGCGCTGCGCTCCCGCGGCTTGAGGAGACATCGGCGCTGAGTCCCG	110
Db	1	GCCCTGAGAGCCCGGCGCTGCGCTCCCGCGGCTTGAGGAGACATCGGCGCTGAGTCCCG	60
QY	111	TCCCGAGGCTGAGGCTGCGGGGACCGCGCCACCGCTGCGCCAGAGAGTGGGGCTCTCTGC	170
Db	61	TCCCGAGGCTGAGGCTGCGGGGACCGCGCCACCGCTGCGCCAGAGAGTGGGGCTCTCTGC	120
QY	171	CTTCTTCCAGGCCCCACGTTGCTGGCCGCTTGCGCGAGTGGGCCGCACTGCTCTGCT	230
Db	121	CTTCTTCCAGGCCCCACGTTGCTGGCCGCTTGCGCGAGTGGGCCGCACTGCTCTGCT	180
QY	231	TGGGCGACCTCTGCGCCCCCGGCGCTTGCGCTGAGGGGCTCTGAGTCTCTTGGGGCTC	290
Db	181	TGGGCGACCTCTGCGCCCCCGGCGCTTGCGCTGAGGGGCTCTGAGTCTCTTGGGGCTC	240
QY	291	CTGGGAGCATCGACGCCCGGAGCGGTCCTCATATAGCTCGGAGAACAGACTGCAGG	350
Db	241	CTGGGAGCATCGACGCCCGGAGCGGTCCTCATATAGCTCGGAGAACAGACTGCAGG	300
QY	351	GACCGAGAAAAGAAATATGATGAGCCCCAGCACCGCATCTGCTGCTCCCGCTGCCGCA	410
Db	301	GACCGAGAAAAGAAATATGATGAGCCCCAGCACCGCATCTGCTGCTCCCGCTGCCGCA	360
QY	411	GGCAGCTATGTCACGCTAATATGATGAGCCGATCCGGGACACAGTTTGGCCACATGTGCC	470

Db	361	GGCACCTATGTTCTCAGCTAAATGTAATGTAACCGCATTCGGGAAACAGATTGTGGCAACAATGTGGC	420
QY	471	GAGATTTCTTACAAGCAGCAGCTGGAACTTACCTGACCATCTGCGAGCTGTGCGCCCTGT	530
Db	421	GAGAAATTCCTACAAGCAGCAGCTGGAACTTACCTGACCATCTGCGAGCTGTGCGCCCTGT	480
QY	531	GACCAATGATGAGGGCTCCGAGAGGATTTGGCCCCCTGCACAGCAAGCAAGGAAAGCCAGGCG	590
Db	481	GACCAATGATGAGGGCTCCGAGAGGATTTGGCCCCCTGCACAGCAAGCAAGGAAAGCCAGGCG	540
QY	591	CGCTGCAGCGGGGAAATGTTCTGTGCTGGCTGGGCCCTTGATGTATCACTGTGGAGCTA	650
Db	541	CGCTGCAGCGGGGAAATGTTCTGTGCTGGCTGGGCCCTTGATGTATCACTGTGGAGCTA	600
QY	651	CTTTCTGACTGCCCCGCTGGCACTGAAAGCCGAGCTCAAAAGATGAAGTTGGGAAGGTTAAC	710
Db	601	CTTTCTGACTGCCCCGCTGGCACTGAAAGCCGAGCTCAAAAGATGAAGTTGGGAAGGTTAAC	660
QY	711	AACCACTGGCTCCCTGCGAAGGCGAGGGGCACTTCGAAATATCCTCCTCCCGCAGCGCCGC	770
Db	661	AACCACTGGCTCCCTGCGAAGGCGAGGGGCACTTCGAAATATCCTCCTCCCGCAGCGCCGC	720
QY	771	TGCAGCGCCACACACAGGTGTGAGAACCAAGGTCTGTGTGAGAGCGACTCCAGGCACTGCC	830
Db	721	TGCAGCGCCACACACAGGTGTGAGAACCAAGGTCTGTGTGAGAGCGACTCCAGGCACTGCC	780
QY	831	CAGTTCGACACAAACCTGCAAAAAATTCATTAGAGCCACTGCCCCAGAGATGTCA-----	884
Db	781	CAGTTCGACACAAACCTGCAAAAAATTCATTAGAGCCACTGCCCCAGAGATGTCAAGAAC	840
QY	885	-----	884
Db	841	ATGCGTATGCTGGGCCGTTCTGCTGCCACTGGCCCTTCTTGTGCTTGCACACGCTTTC	900
QY	885	-----	902
Db	901	TCCTGCACTGTGGAAAGACCACCTTCTCTCTGAGGAAACTGGGATCGCTGTCAAGAGG	960
QY	903	CGTCCGACGAGGAGAGGAGACCCAAATCTCTGTAAGCTGGAAGCTGGGAGCTCCGGAAGGCCAT	962
Db	961	CGTCCGACGAGGAGAGGAGACCCAAATCTCTGTAAGCTGGAAGCTGGGAGCTCCGGAAGGCCAT	1020
QY	963	CCATACTTCCCTGACTGTGTGATACAGCCACTGTACCCATTCTGAGAGATGTTTCCCACTA	1080
Db	1021	CCATACTTCCCTGACTGTGTGATACAGCCACTGTGTACCCATTCTGAGAGATGTTTCCCACTA	1080
QY	1023	TTCGACTGGGGTCCCCGGAGGCCAGTTTGGAGGGCAGGGGTGCCGAACAGCAGAGTCTT	1082
Db	1081	TTCGACTGGGGTCCCCGGAGGCCAGTTTGGAGGGCAGGGGTGCCGAACAGCAGAGTCTT	1140
QY	1083	CTGACCTGACACAGGAGCGCAGTTTGAACCCGGGGAGCAGAGCCAGGTGGCCCAAGGT	1142
Db	1141	CTGAGCTGACACAGGAGCGCAGTTTGAACCCGGGGAGCAGAGCCAGGTGGCCCAAGGT	1200
QY	1143	ACCAATGGCACTTATGTACACGGCGGGTCTATGATCTATACGTGGCAATCTACTATCTAC	1202
Db	1201	ACCAATGGCACTTATGTACACGGCGGGTCTATGATCTATACGTGGCAATCTACTATCTAC	1260
QY	1203	AATGACACAGTACTGGGGGGAGACCGGGGTCTGAGAGCTCTCCAGAGTACCCCGCAACT	1262
Db	1261	AATGACACAGTACTGGGGGGAGACCGGGGTCTGAGAGCTCTCCAGAGTACCCCGCAACT	1320
QY	1263	CCATATCCCATTTCCGGAAGAGGGGGAACCTTGAGCCCTCCGGGCTCTCTACACCCCAACAG	1322
Db	1321	CCATATCCCATTTCCGGAAGAGGGGGAACCTTGAGCCCTCCGGGCTCTCTACACCCCAACAG	1380
QY	1323	GAAATGGCAAGGCTTGGCACTTACGCGAGACAGACACTGTGTGTCCACACCTCTTAACT	1382
Db	1381	GAAATGGCAAGGCTTGGCACTTACGCGAGACAGACACTGTGTGTCCACACCTCTTAACT	1440
QY	1383	AGGGGCCCAAGAACCAATTTATCACCCCATGACTGACTGATGTAGAAAAGCGGAAGA	1442

US-09-962-436-262
; Sequence 262, Application US/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
; APPLICANT: Soppel, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689230-75
; CURRENT APPLICATION NUMBER: US/09-962,436
; PRIOR APPLICATION NUMBER: US/60/235,082
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/234,924
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 262
; LENGTH: 2136
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-436-262

Query Match 91.3%; Score 1809.2; DB 10; Length 2136;
Best Local Similarity 94.6%; Pred. No. 0;
Matches 1929; Conservative 0; Mismatches 3; Indels 108; Gaps 1;

Oy 51 GCCCTGGAGGCGCGGCTGCGGCGCTGCGGCGCTGCGGCGTGCACATCGGCGCTGAGTCCG 110
Db 1 GCCCTGGAGGCGCGGCTGCGGCGCTGCGGCGCTGCGGCGTGCACATCGGCGCTGAGTCCG 60
Oy 111 TCCAGGCTTGGGCTCGGGGAGCGCGCGCCACCGCTGCCAGAGAGCTGGGGCTCTGC 170
Db 61 TCCAGGCTTGGGCTCGGGGAGCGCGCGCCACCGCTGCCAGAGAGCTGGGGCTCTGC 120
Oy 171 CTCTCTCCAGGCGCGGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCT 230
Db 121 CTCTCTCCAGGCGCGGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCT 180
Oy 231 TGGGCGACCTGCGCGCGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCT 290
Db 181 TGGGCGACCTGCGCGCGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCT 240
Oy 291 CTGGCAGCATGCGAGCGCGCGCGGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGG 350
Db 241 CTGGCAGCATGCGAGCGCGCGCGGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGG 300
Oy 351 GACCGAGAAAGGAACTATGAGCGCGCGAGCGCGCATCTGCTGCTGCGGCGCGCGCA 410
Db 301 GACCGAGAAAGGAACTATGAGCGCGCGAGCGCGCATCTGCTGCTGCGGCGCGCGCA 360
Oy 411 GGCACCTATGCTCAGCTAAATGTAGCCGATCCGGGACACAGTTTGGCCATGTC 470
Db 361 GGCACCTATGCTCAGCTAAATGTAGCCGATCCGGGACACAGTTTGGCCATGTC 420
Oy 471 GAGAAATCTTACACGAGCACTGGAATCTGACCATCTGCGGCGCGCGCTG 530
Db 421 GAGAAATCTTACACGAGCACTGGAATCTGACCATCTGCGGCGCGCGCTG 480
Oy 531 GACCCAGTATGAGGCGCTGAGGAGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 590
Db 481 GACCCAGTATGAGGCGCTGAGGAGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
Oy 591 GCGTCCAGCGCGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 650
Db 541 GCGTCCAGCGCGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Oy 651 CTTTCTGACTGCGCGCGCTGCGGCTGAGGCGAGCTCAAGATGAGAGTTGGAGGTAAC 710
Db 601 CTTTCTGACTGCGCGCGCTGCGGCTGAGGCGAGCTCAAGATGAGAGTTGGAGGTAAC 660
Oy 711 AACCACTGAGTCCCTGGAAGGAGGAGCTTCAGAAATACCTCTCCCGCGCGCGCGCG 770
Db 661 AACCACTGAGTCCCTGGAAGGAGGAGCTTCAGAAATACCTCTCCCGCGCGCGCGCGCG 720

Oy 771 TGCCAGCCCAACACAGGTGTGAGAACCAAGTCTGTGAGGAGGCTCCAGGCACTGCC 830
Db 721 TGCCAGCCCAACACAGGTGTGAGAACCAAGTCTGTGAGGAGGCTCCAGGCACTGCC 780
Oy 831 CAGTCCGACACACCTGCAAAAAATTCATTAGAGCCACTGCCCCAGAGATGTCA----- 884
Db 781 CAGTCCGACACACCTGCAAAAAATTCATTAGAGCCACTGCCCCAGAGATGTCAAGAAC 840
Oy 885 ----- 884
Db 841 ATGCTATGCTGCCGTTCTGCTGCGCACTGCGCTTCTTCTGCTCTCCACCGCTTTC 900
Oy 885 -----GATCGCTGCTCAAGAG 902
Db 901 TCCATGATCTGGAGAGGAGCCACCTTCTCTGCAAGAACTGGGATGCTGCTCAAGAG 960
Oy 903 GGTCCGAGGAGAGAGGAGCCCAATCTGTAGCTGGAAGCTGGAGCCCTCCAGAGCCCAT 962
Db 961 GGTCCGAGGAGAGAGGAGCCCAATCTGTAGCTGGAAGCTGGAGCCCTCCAGAGCCCAT 1020
Oy 963 CCATACCTTCCCTGACTTGTGTACAGGCACTGCTACCCATTCTGTGAGATGTTTCCCACTA 1022
Db 1021 CCATACCTTCCCTGACTTGTGTACAGGCACTGCTACCCATTCTGTGAGATGTTTCCCACTA 1080
Oy 1023 TCCACTGCGCTCCCGCGAGCCCAATTTTGGAGGAGGCGGTGCCCAACAGCAGAGTCT 1082
Db 1081 TCCACTGCGCTCCCGCGAGCCCAATTTTGGAGGAGGCGGTGCCCAACAGCAGAGTCT 1140
Oy 1083 CTGACCTGACAGGAGAGCCGAGTTGGAAACCGGGGAGCAGAGCAGAGTGGCCACG 1142
Db 1141 CTGACCTGACAGGAGAGCCGAGTTGGAAACCGGGGAGCAGAGCAGAGTGGCCACG 1200
Oy 1143 ACCAATGGCATTCATGTACCGCGCGGTCTATGACTATACCTGGAACATCTACATCTAC 1202
Db 1201 ACCAATGGCATTCATGTACCGCGCGGTCTATGACTATACCTGGAACATCTACATCTAC 1260
Oy 1203 AATGACACGATCTGAGGAGGAGCACCAGGCTCTGAGAGCTCCAGCTACCCCGAACT 1262
Db 1261 AATGACACGATCTGAGGAGGAGCACCAGGCTCTGAGAGCTCCAGCTACCCCGAACT 1320
Oy 1263 CCATACCCCATTTCCGAAGAGGAGGAGCTGAGGCTCCCGGGCTCTTACACCCACAG 1322
Db 1321 CCATACCCCATTTCCGAAGAGGAGGAGCTGAGGCTCCCGGGCTCTTACACCCACAG 1380
Oy 1323 GAGATGGCAAGGCTTGGCACCTAGCAGAGAGAGCACTGAGGTGGGCCACACCTCTAAC 1382
Db 1381 GAGATGGCAAGGCTTGGCACCTAGCAGAGAGAGCACTGAGGTGGGCCACACCTCTAAC 1440
Oy 1383 AGGGGCCCAAGGAACCAATTTATCACCCCATGACTGAGTCTGAGAAAAAGGAGAA 1442
Db 1441 AGGGGCCCAAGGAACCAATTTATCACCCCATGACTGAGTCTGAGAAAAAGGAGAA 1500
Oy 1443 AGGGGGCACAAGGCACTTCTGCTTGAAGGCTGCTGCGCCACAGTGGGATTCAGAGG 1502
Db 1501 AGGGGGCACAAGGCACTTCTGCTTGAAGGCTGCTGCGCCACAGTGGGATTCAGAGG 1560
Oy 1503 GCGTATGAGGCGCGGGGAGAGAGCCCTAAGGATTAAGGCTCAGACACCTGTGAGA 1562
Db 1561 GCGTATGAGGCGCGGGGAGAGAGCCCTAAGGATTAAGGCTCAGACACCTGTGAGA 1620
Oy 1563 GCAGGTGGGCACTGCTGGGTACGCTGCTCCACAGGAGCTTCCCTACTGCTAGAGCA 1622
Db 1621 GCAGGTGGGCACTGCTGGGTACGCTGCTCCACAGGAGCTTCCCTACTGCTAGAGCA 1680
Oy 1623 ACCTGAGGCTTCCCGGAGAGCCACACCCCTGAGGCTGCTGAGCTCAGGCGAGGAG 1682
Db 1681 ACCTGAGGCTTCCCGGAGAGCCACACCCCTGAGGCTGCTGAGCTCAGGCGAGGAG 1740
Oy 1683 AGGCGACATGATACCAATCTGCTGCTTACAGCAGCGCGGACCGGAGCAGGCGAG 1742
Db 1741 AGGCGACATGATACCAATCTGCTGCTTACAGCAGCGCGGACCGGAGCAGGCGAG 1800

TYPE: DNA
ORGANISM: Homo sapiens
US-09-954-531-1348

Query Match 91.3%; Score 1809.2; DB 9; Length 2136;
Best Local Similarity 94.6%; Pred. No. 0;
Matches 1929; Conservative 0; Mismatches 3; Indels 108; Gaps 1;

51 GCCCTGAGGCGCCGCTGCGCCCTCCGCGCTGGGGTGACATCGGCCCTAGTCCG 110
1 GCCCTGAGGCGCCGCTGCGCCCTCCGCGCTGGGGTGACATCGGCCCTAGTCCG 60
111 TCCAGGCTTGGGCTTGGGCGACCGCGCCAGCGCTGCCAGAGCTGGGGCTCTGC 170
61 TCCAGGCTTGGGCTTGGGCGACCGCGCCAGCGCTGCCAGAGCTGGGGCTCTGC 120
171 CTTCCTCCAGGCGCCGCTGCGCCCTCCGCGCTGGGGTGACATCGGCCCTAGTCC 230
121 CTTCCTCCAGGCGCCGCTGCGCCCTCCGCGCTGGGGTGACATCGGCCCTAGTCC 180
231 TGGGCGACCTCTGCGCCGCTGCGCCCTCCGCGCTGGGGTGACATCGGCCCTAG 290
181 TGGGCGACCTCTGCGCCGCTGCGCCCTCCGCGCTGGGGTGACATCGGCCCTAG 240
291 CTGGGACATCGGCGCCGCGCTGCGCCCTCCGCGCTGGGGTGACATCGGCCCTAG 350
241 CTGGGACATCGGCGCCGCGCTGCGCCCTCCGCGCTGGGGTGACATCGGCCCTAG 300
351 GACGAGAAAAAGAAATCTATGAGCCCGACACCGCATCTGCTCCGCTGCCGCA 410
301 GACGAGAAAAAGAAATCTATGAGCCCGACACCGCATCTGCTCCGCTGCCGCA 360
411 GGACCATCTGTCAGCTAATATGAGCGCATCCGCGACAGTTTGGCAATGAGCC 470
361 GGACCATCTGTCAGCTAATATGAGCGCATCCGCGACAGTTTGGCAATGAGCC 420
471 GAGAAATCTTACAAGAGCACTGGAATACCTGACATCTGCGCCAGTGTGCCCTGT 530
421 GAGAAATCTTACAAGAGCACTGGAATACCTGACATCTGCGCCAGTGTGCCCTGT 480
531 GACCAATGATGGGCTTCGAGAGATTGCCCTGCAACAGCAAGGAAAGCCAGTGC 590
481 GACCAATGATGGGCTTCGAGAGATTGCCCTGCAACAGCAAGGAAAGCCAGTGC 540
591 CGCTGGCAGCGGGAATGTTGTCGCTGCGCCCTCCGCGCTGGGGTGACATCGGAGTA 650
541 CGCTGGCAGCGGGAATGTTGTCGCTGCGCCCTCCGCGCTGGGGTGACATCGGAGTA 600
651 CTTCCTGACCTGCGCCCTGCGCACTGAAGCGAGCTCAAGATGAAGTTGGGAAGGTAC 710
601 CTTCCTGACCTGCGCCCTGCGCACTGAAGCGAGCTCAAGATGAAGTTGGGAAGGTAC 660
711 AACCACTGCGCTGCGCTGCGCACTGAAGCGAGCTCAAGATGACTCTCCCGAGCGCCG 770
661 AACCACTGCGCTGCGCTGCGCACTGAAGCGAGCTCAAGATGACTCTCCCGAGCGCCG 720
771 TGGCAGCGCCGACAGCTGAGAACCAAGCTGAGGAGGAGCTCCAGGAGTGC 830
721 TGGCAGCGCCGACAGCTGAGAACCAAGCTGAGGAGGAGCTCCAGGAGTGC 780
831 CAGTCCGACACACCTGCAAAATCCATTAGAGCCACTGCCCGAGAGATGCA----- 884
781 CAGTCCGACACACCTGCAAAATCCATTAGAGCCACTGCCCGAGAGATGCAAGAAC 840
885 ----- 884
841 ATGTGATGCTGGCGCTTCTGCTGCCACTGGCCTTCTTCTGCTCTTGGCAACGCTTC 900
885 -----GGATCGCTGCTCAAGAG 902
901 TCTGCACTGTGAGAGGAGCACCCTTCTCTGCAAGAAACCTGGGATGCTCTCAAGAG 960
903 CGTCCGAGGAGAGGAGCACCCTTCTGTAAGTGGAGCTGGAGCTCCGAGAGCCCAT 962

661 CGTCCGAGGAGAGGAGCACCCTTCTGTAAGTGGAGCTGGAGCTCCGAGAGCCCAT 1020
961 CCATACCTCCCTGACTTGGTACAGCCAGCTACCCATTCTTGAGATGTTTCCCACTA 1022
1021 CCATACCTCCCTGACTTGGTACAGCCAGCTACCCATTCTTGAGATGTTTCCCACTA 1080
1023 TCCACTGGGCTCCCGCAGCCCGAGTTTGGAGGCGGGGGTGGCGCAAGAGAGTCC 1082
1081 TCCACTGGGCTCCCGCAGCCCGAGTTTGGAGGCGGGGGTGGCGCAAGAGAGTCC 1140
1083 CTGACCTGACACAGGAGCGAGCTGTTGAAACCGGAGAGAGCCAGGTGGCCCACT 1142
1141 CTGACCTGACACAGGAGCGAGCTGTTGAAACCGGAGAGAGCCAGGTGGCCCACT 1200
1143 ACCAATGGCATTCATCTACACCGCGGCTCTGACTATCTACTGCAACATCTACTAC 1202
1201 ACCAATGGCATTCATCTACACCGCGGCTCTGACTATCTACTGCAACATCTACTAC 1260
1203 AATGACACGACTTGGGGGAGACCGGGTCTGAGAGACCTCCAGCTACCCCGAAGCT 1262
1261 AATGACACGACTTGGGGGAGACCGGGTCTGAGAGACCTCCAGCTACCCCGAAGCT 1320
1263 CCATACCCCATTTCCGAGAGGGGGAACCTGCGCTCCCGGCTCTACTACCCCAACAG 1322
1321 CCATACCCCATTTCCGAGAGGGGGAACCTGCGCTCCCGGCTCTACTACCCCAACAG 1380
1323 GAAGATGGCAAGCTTGGCACTAGGGGAGACAGACACTGTGTGTCACACCTCTAAC 1382
1381 GAAGATGGCAAGCTTGGCACTAGGGGAGACAGACACTGTGTGTCACACCTCTAAC 1440
1383 AGGGGCCCAAGAAACATTTATCACCATGACTGACTGAGTGTGAGAAAGGAGAGAG 1442
1441 AGGGGCCCAAGAAACATTTATCACCATGACTGAGTGTGAGAAAGGAGAGAG 1500
1443 AGGGGGGCAAGAGGCACTTTCCTTGAAGGCTGCCCTGCCAGTGGGATTCACAGGG 1502
1501 AGGGGGGCAAGAGGCACTTTCCTTGAAGGCTGCCCTGCCAGTGGGATTCACAGGG 1560
1503 GCGTGAATGAGGGCCCGGAGAGCAGAGCCCTAAGAGGATTAAGGCTCAGACCTGTAGA 1562
1561 GCGTGAATGAGGGCCCGGAGAGCAGAGCCCTAAGAGGATTAAGGCTCAGACCTGTAGA 1620
1621 GAGGTGGCACTGGGCTGGGTAGGTCCTCCACAGAGACTTCCCTACTGCTGAGCAA 1680
1623 ACCTGAGGCTCCCGGAGAGCACCACCCCTGGGGGCTGCTAGGCTCAGAGGAG 1682
1681 ACCTGAGGCTCCCGGAGAGCACCACCCCTGGGGGCTGCTAGGCTCAGAGGAG 1740
1683 AGGCGCATGATACCACTCTGCGCACTAGACAGCCGACAGCGAGAGCAGCAGAG 1742
1741 AGGCGCATGATACCACTCTGCGCACTAGAGGAGCGGACAGCGAGAGCAGCAGAG 1800
1743 GAGCGCGCACAGGCTACCTGCAAGAGCTCAAGGCGCCCTCTTAAGATTTGTGTGC 1802
1801 GAGCGCGCACAGGCTACCTGCAAGAGCTCAAGGCGCCCTCTTAAGATTTGTGTGC 1860
1803 TCAATCCCAAGCTCAGAGACCTTGGGGTTCACACTCAAGTGAAGTGAAGTGAAGC 1862
1861 TCAATCCCAAGCTCAGAGACCTTGGGGTTCACACTCAAGTGAAGTGAAGTGAAGC 1920
1863 CTGATGAAGATGAATTTATAGGAGAGCCTCTTCCCTCCCTCTAGAGAGAGAA 1922
1921 CTGATGAAGATGAATTTATAGGAGAGCAGCCTCTTCCCTCTAGAGAGAGAA 1980
1923 AGGAGTCAATTAACACTAGGGGCTTGGGTAGATTTCTTAGATTAAGGAGAGATTTGG 1982
1981 AGGAGTCAATTAACACTAGGGGCTTGGGTAGATTTCTTAGATTAAGGAGAGATTTGG 2040

RESULT 3

QY 241 CTGCCCCGGGCTGGCTGGGGGCTCTGTGCTGGGCTCTTGGGGCTCTGGGACAT 300
| | | | |
Db 241 CTGCCCCGGGCTGGCTGGGGGCTCTGTGCTGGGCTCTTGGGGCTCTGGGACAT 300
QY 301 CGCAGCCCCAGGGGCTGCTCATATGCTGGAGAACAGACTGCGAGGACAGGAAA 360
| | | | |
Db 301 CGCAGCCCCAGGGGCTGCTCATATGCTGGAGAACAGACTGCGAGGACAGGAAA 360
QY 361 AGGAATCTATGAGCCCGACAGCCGATCTGTCTCCGCGGGCCGACGACCTATG 420
| | | | |
Db 361 AGGAATCTATGAGCCCGACAGCCGATCTGTCTCCGCGGGCCGACGACCTATG 420
QY 421 TCTGAGTAAATGTAGCCGCATCCCGGACACAGTTTGTGCCACATGTCCGAGATTCT 480
| | | | |
Db 421 TCTGAGTAAATGTAGCCGCATCCCGGACACAGTTTGTGCCACATGTCCGAGATTCT 480
QY 481 ACAAGAGACACTGGAATCACTGACCAATCTGCGACGCTGTGCGCCCTGTGACCCAGTA 540
| | | | |
Db 481 ACAAGAGACACTGGAATCACTGACCAATCTGCGACGCTGTGCGCCCTGTGACCCAGTA 540
QY 541 TGGGCTGGAGGAGATTCGCCCCCTGCACAGAAACGGAACCGAGTGGCGCTGCAGC 600
| | | | |
Db 541 TGGGCTGGAGGAGATTCGCCCCCTGCACAGAAACGGAACCGAGTGGCGCTGCAGC 600
QY 601 CGGGAATGTCTGTGCTGCTGCTGGGCTGCGAGTGTACACACTGCGAGTACTTCTGACT 660
| | | | |
Db 601 CGGGAATGTCTGTGCTGCTGCTGGGCTGCGAGTGTACACACTGCGAGTACTTCTGACT 660
QY 661 GCGGCGCTGGCAGCTAGAGCCGAGCTCAAGATGAAGTTGGAGGGTAAACAACACTGCG 720
| | | | |
Db 661 GCGGCGCTGGCAGCTAGAGCCGAGCTCAAGATGAAGTTGGAGGGTAAACAACACTGCG 720
QY 721 TCCCTGGAAGGAGGAGGAGCTTCAGAAATCTCTCCCGGCGGCTGGCCGACGCC 780
| | | | |
Db 721 TCCCTGGAAGGAGGAGGAGCTTCAGAAATCTCTCCCGGCGGCTGGCCGACGCC 780
QY 781 ACACCAAGGTGTGAGAACCAAGGTGTGTGAGAGGAGCTCCAGGCACTGCGAGTCCGACA 840
| | | | |
Db 781 ACACCAAGGTGTGAGAACCAAGGTGTGTGAGAGGAGCTCCAGGCACTGCGAGTCCGACA 840
QY 841 CAACCTGCAAAAATCCATTAGAGCACTGCCCCCAGAGATGTGAGAGTCTGCTCAAGA 900
| | | | |
Db 841 CAACCTGCAAAAATCCATTAGAGCACTGCCCCCAGAGATGTGAGAGTCTGCTCAAGA 900
QY 901 GGGCTGCCGAGGAGGAGGAGCCCAATCTGTAGCTGGAAGCTGGAGAGCTCCGAGAGGCC 960
| | | | |
Db 901 GGGCTGCCGAGGAGGAGGAGCCCAATCTGTAGCTGGAAGCTGGAGAGCTCCGAGAGGCC 960
QY 961 ATCATACTTCCCTGACTTGTGTAGACCACTGCTACCCATTCTGTGAGATGTTTCCGAG 1020
| | | | |
Db 961 ATCATACTTCCCTGACTTGTGTAGACCACTGCTACCCATTCTGTGAGATGTTTCCGAG 1020
QY 1021 TATTCACCTGGGCTCCCCGAGCCGAGTTTGGAGGAGGGGTGCCCAACAGCAGATC 1080
| | | | |
Db 1021 TATTCACCTGGGCTCCCCGAGCCGAGTTTGGAGGAGGGGTGCCCAACAGCAGATC 1080
QY 1081 CTTCTGAGCTGACGAGGAGCCGAGTTGGAACCCGGGAGAGAGCCAGGGTGGCCACG 1140
| | | | |
Db 1081 CTTCTGAGCTGACGAGGAGCCGAGTTGGAACCCGGGAGAGAGCCAGGGTGGCCACG 1140
QY 1141 GTACCAATGGCAATTCATGTCACCGGGGTATGACTATCACTGCAACATCTACATCT 1200
| | | | |
Db 1141 GTACCAATGGCAATTCATGTCACCGGGGTATGACTATCACTGCAACATCTACATCT 1200
QY 1201 ACAATGAGACACTGAGGGGAGACACCGGGTCTGTGAGAGCTCCAGTACCCCGAAC 1260
| | | | |
Db 1201 ACAATGAGACACTGAGGGGAGACACCGGGTCTGTGAGAGCTCCAGTACCCCGAAC 1260
QY 1261 CTCCTATACCCCTATTCGGAAGGGGAGACCTGGGCTCCGGGGCTCTCTACACCCACC 1320
| | | | |
Db 1261 CTCCTATACCCCTATTCGGAAGGGGAGACCTGGGCTCCGGGGCTCTCTACACCCACC 1320
QY 1321 AGGAAGATGGCAAGGCTTGGACCTAGCGGAGACAGACATGTGTGTCACACCTCTTA 1380
| | | | |

Db 1321 AGGAAGATGGCAAGGCTTGGACACCTAGCGGAGACAGACACTGTGTGCTCCACCTCTA 1380
QY 1381 ACAGGGGCCCAAGAGCAACCAATTTATACACCCTAGACTGACTGATGTGAGAAAAGGCAAA 1440
| | | | |
Db 1381 ACAGGGGCCCAAGAGCAACCAATTTATACACCCTAGACTGACTGATGTGAGAAAAGGCAAA 1440
QY 1441 GAAGGGGGGCAAGGGGACCTTCTCCCTTGAAGGCTGCGCTGCCAGCTGGGATTCACAG 1500
| | | | |
Db 1441 GAAGGGGGGCAAGGGGACCTTCTCCCTTGAAGGCTGCGCTGCCAGCTGGGATTCACAG 1500
QY 1501 GGGCTGAGTAGAGGGCCCGGGGAGCAGACCCCTAAGGGATTAAAGGCTCAGACACTCTGA 1560
| | | | |
Db 1501 GGGCTGAGTAGAGGGCCCGGGGAGCAGAGCCCTAAGGGATTAAAGGCTCAGACACTCTGA 1560
QY 1561 GAGCAGGTGGGCACTGCTGGGTACGAGTGTGCTCCACAGAGACTCTCCCTACTGCTGAGC 1620
| | | | |
Db 1561 GAGCAGGTGGGCACTGCTGGGTACGAGTGTGCTCCACAGAGACTCTCCCTACTGCTGAGC 1620
QY 1621 AAACCTGAGGGCTCCGGGACAGCCACCCAGCCCTGGGGGCTGCTCAGCTCAGGACAGG 1680
| | | | |
Db 1621 AAACCTGAGGGCTCCGGGACAGCCACCCAGCCCTGGGGGCTGCTCAGCTCAGGACAGG 1680
QY 1681 ACAGGGCACTGATACCAACTGCTGCTCCCACTACAGACGCGCACCGGAGCAGCAGCG 1740
| | | | |
Db 1681 ACAGGGCACTGATACCAACTGCTGCTCCCACTACAGACGCGCACCGGAGCAGCAGCG 1740
QY 1741 AGGAGCCGCCACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
| | | | |
Db 1741 AGGAGCCGCCACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
QY 1801 GCTCATCCCAAGCTTCAAGAGCCCTTGGGGTTCACACTTACAGTGGAGTGGAGTGA 1860
| | | | |
Db 1801 GCTCATCCCAAGCTTCAAGAGCCCTTGGGGTTCACACTTACAGTGGAGTGGAGTGA 1860
QY 1861 CCTCATGAGATGAATTAATATAGGAGAGAGCTCTCCCTCCCTCTAGAGAGAGG 1920
| | | | |
Db 1861 CCTCATGAGATGAATTAATATAGGAGAGAGAGCTCTCCCTCCCTCTAGAGAGAGG 1920
QY 1921 AAAGGAGCTATTAAACAATAGGGGTTGGGTAGGATTCATAGTATGGGAAAGTTT 1980
| | | | |
Db 1921 AAAGGAGCTATTAAACAATAGGGGTTGGGTAGGATTCATAGTATGGGAAAGTTT 1980
QY 1981 GG 1982
| | | | |
Db 1981 GG 1982
| | | | |
RESULT 2
US-09-954-531-1348
; Sequence 1348, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
; FILE REFERENCE: 689290-77
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,133
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1348
; LENGTH: 2136

QY	1	GCCTCGGAGCCACGCTCGGTCCACTCCACTCCTGAGCTCCGCTCATGGAGAGCCCTGAGAG	60
Db	1	GCCTCGGAGCCACGCTCGGTCCACTCCACTCCTGAGCTCCGCTCATGGAGAGCCCTGAGAG	60
QY	61	CCCGGCGTGGCGGCTCCCGGGCCCTGGGGTGCATCCGCGCTGAGTCCGCTCCAGGCTTC	120
Db	61	CCCGGCGTGGCGGCTCCCGGGCCCTGGGGTGCATCCGCGCTGAGTCCGCTCCAGGCTTC	120
QY	121	TGGGCGTCGGGACGCGCGCCACCGCTGCCAGGACGTCGGGCGCTCCCTTCCTCCCA	180
Db	121	TGGGCGTCGGGACGCGCGCCACCGCTGCCAGGACGTCGGGCGCTCCCTTCCTCCCA	180
QY	181	GGCCCCACAGTGTGTCGCGGCTTGCGCGAGTGGCGCCATGCTCTTCCTTGGGCGACCT	240
Db	181	GGCCCCACAGTGTGTCGCGGCTTGCGCGAGTGGCGCCATGCTCTTCCTTGGGCGACCT	240

Query Match	10.7%	Score 239.5	DB 1	Length 401
Best Local Similarity	30.9%	Pred. No. 3e-07		
Matches	51	Conservative	21	Mismatches 84; Indels 9; Gaps 4;

QY	52	EPQHRICSRCPPTGYVAKSCSRIDYQACCAESENSYEHNNYLTICQLCRP-CDPYKGL	110
		: : : : : : : : : : : : : : : : : : : : : : : : :	
Db	34	ETSHQLLDKCPPTGYLLQHCIAKKRYTACAPQPHIYITDSMTSECLCYSCVKELOQY	93
		: : : : : : : : : : : : : : : : : : : : : : : : :	
QY	111	EEIAPCTSKRTQCCQPGMFCAMALDCTHCELLSDCPGTAEALKDEYGGNNHCYPC	170
		: : : : : : : : : : : : : : : : : : : : : : : : :	
Db	94	KQ--ECNRTIHNRYVCECKRGRY----LEIERFLKHSRSCPGR-GVVOAGTPERNTVYCKRC	145
		: : : : : : : : : : : : : : : : : : : : : : : : :	
QY	171	KAGHFOHTSSPAPCRPHTRCENGLVAAAPTAQSDTCKNPLE	215
		: : : : : : : : : : : : : : : : : : : : : : : : :	
Db	146	PDGFSNMTSSKAPCRKHTTNCVFEILLTQKGNAIHDNICSNSE	190
		: : : : : : : : : : : : : : : : : : : : : : : : :	

Search completed: April 8, 2003, 10:49:40
Job time : 30.6158 secs

THIS PAGE BLANK (USPTO)

RC TISSUE-Lung cancer;
 RX MEDLINE=98151033; PubMed=9492069;
 RA Yasuda H., Shima N., Nakagawa N., Mochizuki S.-I., Yano K., Fujise N.,
 RA Sato Y., Goto M., Yamaguchi K., Kuriyama M., Kanno T., Murakami A.,
 RA Tsuda E., Morinaga T., Higashio K.;
 RT "Identity of osteoclastogenesis inhibitory factor (OCIF) and
 RT osteoprotegerin (OPG): a mechanism by which OPG/OCIF inhibits
 RT osteoclastogenesis in vitro.";
 RL Endocrinology 139:1329-1337(1998).
 RP [3]
 RP SEQUENCE FROM N.A., AND VARIANT ASN-3.
 RC TISSUE-Placenta;
 RX MEDLINE=9831569; PubMed=9688283;
 RA Morinaga T., Nakagawa N., Yasuda H., Tsuda E., Higashio K.;
 RT "Cloning and characterization of the gene encoding human
 RT osteoprotegerin/osteoclastogenesis-inhibitory factor.";
 RL Eur. J. Biochem. 254:685-691(1998).
 RN [4]
 RN SEQUENCE FROM N.A., AND VARIANT ASN-3.
 RC TISSUE-Eye;
 RA Strausberg R.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RN SEQUENCE OF 22-36 AND 378-401.
 RX MEDLINE=98238645; PubMed=9571159;
 RA Tomoyasu A., Goto M., Fujise N., Mochizuki S.-I., Yasuda H.,
 RA Morinaga T., Tsuda E., Higashio K.;
 RT "Characterization of monomeric and homodimeric forms of
 RT osteoclastogenesis inhibitory factor.";
 RL Biochem. Biophys. Res. Commun. 245:382-387(1998).
 RP [6]
 RP SEQUENCE OF 22-393 FROM N.A.
 RC TISSUE-Placenta;
 RA He Z.-Y., Yang G.-Z., Zhang W.-J., Wu X.-F.;
 RT "Cloning and expression of osteoprotegerin from Homo sapiens.";
 RL Acta Biochim. Biophys. Sin. 31:680-684(1999).
 RN [7]
 RN SEQUENCE OF 242-255, 354-359 AND 369-378, AND FUNCTION.
 RX MEDLINE=97312536; PubMed=9168977;
 RA Tsuda E., Goto M., Mochizuki S.-I., Yano K., Kobayashi F.,
 RA Morinaga T., Higashio K.;
 RT "Isolation of a novel cytokine from human fibroblasts that
 RT specifically inhibits osteoclastogenesis.";
 RL Biochem. Biophys. Res. Commun. 234:137-142(1997).
 RN [8]
 RN TRAIL BINDING.
 RP MEDLINE=98269100; PubMed=9603945;
 RX Emery J.G., McDonnell P., Burke M.B., Deen K.C., Lyn S., Silverman C.,
 RA Dul E., Appelbaum E.R., Eichman C., DiPrinzio R., Dadds R.A.,
 RA James I.E., Rosenberg M., Lee J.C., Young P.R.;
 RT "Osteoprotegerin is a receptor for the cytotoxic ligand TRAIL.";
 RL J. Biol. Chem. 273:14363-14367(1998).
 RP [9]
 RP CHARACTERIZATION, AND MUTAGENESIS OF CYS-400.
 RX MEDLINE=98148058; PubMed=9478964;
 RA Yamaguchi K., Kinosaki M., Goto M., Kobayashi F., Tsuda E.,
 RA Morinaga T., Higashio K.;
 RT "Characterization of structural domains of human osteoclastogenesis
 RT inhibitory factor.";
 RL J. Biol. Chem. 273:5117-5123(1998).
 RP [10]
 RP REVIEW.
 RX MEDLINE=21395914; PubMed=11505389;
 RA Hofbauer L.C., Neuber A., Heufelder A.E.;
 RT "Receptor activator of nuclear factor-kappaB ligand and
 RT osteoprotegerin: potential implications for the pathogenesis and
 RT treatment of malignant bone diseases.";
 RL Cancer 92:460-470(2001).
 CC -1- FUNCTION: Acts as decoy receptor for RANKL and thereby neutralizes
 CC its function in osteoclastogenesis. Inhibits the activation of
 CC osteoclasts and promotes osteoclast apoptosis in vitro. Bone
 CC homeostasis seems to depend on the local RANKL/OPG ratio. May also
 CC play a role in preventing arterial calcification. May act as decoy

receptor for TRAIL and protect against apoptosis. TRAIL binding
 blocks the inhibition of osteoclastogenesis.
 -1- SUBUNIT: Homodimer.
 -1- SUBCELLULAR LOCATION: Secreted.
 -1- TISSUE SPECIFICITY: Highly expressed in adult lung, heart, kidney,
 liver, spleen, thymus, prostate, ovary, small intestine, thyroid,
 lymph node, trachea, adrenal gland, testis, and bone marrow.
 Detected at very low levels in brain, placenta and skeletal
 muscle. Highly expressed in fetal kidney, liver and lung.
 -1- INDUCTION: Upregulated by increasing calcium concentration in the
 medium and estrogens. Downregulated by glucocorticoids.
 -1- PTM: N-glycosylated. Contains sialic acid residues.
 -1- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.
 -1- SIMILARITY: CONTAINS 2 DEATH DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: U94332; AAB53709.1; -
 CC EMBL: AB002146; BAA25910.1; -
 CC EMBL: AB008822; BAA32076.1; -
 CC EMBL: AB008821; BAA32076.1; JOINED.
 CC EMBL: BC030155; AAH30155.1; -
 CC EMBL: AF134187; AAF20168.1; -
 CC HSSP: P25942; ICDF.
 CC GeneW: HGNC:11909; TNFRSF11B.
 CC MIM: 602643; -
 CC InterPro: IPR000488; Death.
 CC InterPro: IPR001368; TNFR_c6.
 CC Pfam: PF00020; TNFR_c6; 3.
 CC ProDom: PD000771; TNFR_c6; 1.
 CC SMART: SM00005; DEATH; 4.
 CC SMART: SM00208; TNFR; 1.
 CC PROSITE: PS00017; DEATH DOMAIN; FALSE_NEG.
 CC PROSITE: PS00652; TNFR_NGFR_1; 2.
 CC PROSITE: PS00503; TNFR_NGFR_2; 2.
 CC Receptor; Apoptosis; Glycoprotein; Repeat; Signal; Polymorphism.
 CC SIGNAL
 CC FT 1 21
 CC FT CHAIN 22 401
 CC FT
 CC FT REPEAT 24 62
 CC FT REPEAT 65 105
 CC FT REPEAT 107 142
 CC FT REPEAT 145 185
 CC FT REPEAT 198 269
 CC FT DOMAIN 270 365
 CC FT SITE 400 400
 CC FT
 CC FT DISULFID 41 54
 CC FT DISULFID 44 62
 CC FT DISULFID 65 80
 CC FT DISULFID 83 97
 CC FT DISULFID 87 105
 CC FT DISULFID 107 118
 CC FT DISULFID 124 142
 CC FT DISULFID 145 160
 CC FT DISULFID 166 185
 CC FT CARBOHYD 98 98
 CC FT CARBOHYD 152 152
 CC FT CARBOHYD 165 165
 CC FT CARBOHYD 178 178
 CC FT CARBOHYD 289 289
 CC FT VARIANT 3 3
 CC FT
 CC FT MUTAGEN 400 400
 CC FT MUTAGEN 401 401
 CC FT CONFLICT 263 263
 CC FT SEQUENCE 401 AA; 46040 MW; EDF448B67D86C71E CRC64;

FT	REPEAT	42	75	TNFR-CYS 1.
FT	REPEAT	78	119	TNFR-CYS 2.
FT	REPEAT	121	162	TNFR-CYS 3.
FT	DISULFID	42	53	
FT	DISULFID	54	67	
FT	DISULFID	57	75	
FT	DISULFID	78	93	
FT	DISULFID	96	111	
FT	DISULFID	99	119	
FT	DISULFID	121	138	
FT	DISULFID	127	135	
FT	CARBOHYD	110	110	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	173	173	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARIANT	17	17	K > R.
FT	VARIANT			/FTid=VAR_013007.
FT	VARIANT	241	241	V -> I.
FT	VARIANT			/FTid=VAR_013440.
SO	SEQUENCE	283 AA;	30392 MM;	46CELI3C2C70242C1 CXC64;
	Query Match		10.7%;	Score 240; DB 1; Length 283;
	Best Local Similarity		29.1%;	Pred. No. 2e-07;
	Matches	76;	Conservative	21; Mismatches 96; Indels 68; Gaps 13
QY	10 PGLAMP-----LVLGLEGLLAASQPAPVAPYASENQCROKEKEYEPQHR	56		
Db	4 PG-DWGPPPMWSTPKPTDVLRVLVLYLTFEL-----GACPVALALPSCK--EDEE--PVGS	51		
QY	57 ICSSKCPGGTVVASAKCSRIKDFVCATCAENSYEHMNYLTICOLRCPCDPVMGLEIEIAPC	116		
Db	52 ECCPRCSGGRYRKEMKGELGTGVCEPCPGPIYIAHLNLGSKLOCMCDPMAGLASRNC	111		
QY	117 TSKRRKTQCRCPGMFCAM-ALECTHCELLSDCPPCTEAELKDEVGKGNH----CYPC	170		
Db	112 SRTENAAYCGSCBHGHCIVODGDHCAACRAYATSSPG-----QRVOKGTESQDTLCQNC	165		
QY	171 KAGHQNTSSPSA---RCQPHTRCENOGIVEAPRTAQSDTICKNPDELPPREMGGSLT-	226		
Db	166 PPGTF----SPNGTLEECCOHQTKC-SWLVTAKGAQTSSSHWWM-----FLSGSLVI	212		
QY	227 -----KKRPGE	233		
Db	213 VIVCTVGLLICVKRRKRPRGD	233		
	RESULT 15			
ID	T11B_HUMAN	STANDARD:	PRT:	401 AA.
AC	O00300; O60236; Q9UHP4;			
DC	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Tumor necrosis factor receptor superfamily member 11B precursor			
DE	(Osteoprotegerin) (osteoclastogenesis inhibitory factor).			
GN	TNFRSF1B OR OPG OR OCIF.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
OX	NCBI_TaxId=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Kidney;			
RX	MEMLIN=97262071; PubMed=9108485;			
RA	Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,			
RA	Luetthy R., Nguyen H.Q., Wooden S., Bennett L., Boone T., Shimamoto G.,			
RA	DeRose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,			
RA	Day E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,			
RA	Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,			
RA	Suggs S., Boyle W.J.;			
RT	"Osteoprotegerin: a novel secreted protein involved in the regulation			
RL	of bone density";			
RL	Cell 89:309-319(1997).			
RN	[2]			
RP	SEQUENCE FROM N.A.			

RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Kidney;
 RA Isogai D., Ichino M., Yoshinari M., Yamaura A., Kurokawa F.,
 RA Minami M.;
 RT "Mouse DR6: mouse homolog of human TNFR-related death receptor-6
 (DR6).";
 RT Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Kidney;
 RA Kim V., Machleidt T., Shi W.-X., Wang X., Cai Z.;
 RT "Murine DR6: murine TNFR-related death receptor-6";
 RT Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Strusberg R.;
 RT Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RP FUNCTION.
 RX MEDLINE=21571606; PubMed=11714751;
 RA Zhao H., Yan M., Wang H., Erickson S., Grewal I.S., Dixit V.M.;
 RT "Impaired c-Jun amino terminal kinase activity and T cell
 RT differentiation in death receptor 6-deficient mice.";
 RL J. Exp. Med. 194:1441-1448(2001).
 CC -1- FUNCTION: May activate NF-kappa-B and promote apoptosis (By
 CC similarity). May activate JNK and be involved in T-cell
 CC differentiation.
 CC -1- FUNCTION: May activate NF-kappa-B and promote apoptosis.
 CC May be involved in T-cell differentiation.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
 CC -1- TISSUE SPECIFICITY: Ubiquitous. Highly expressed in adult spleen,
 CC thymus, testis, prostate, ovary, small intestine, colon, brain,
 CC lung and kidney, and in fetal brain, liver and lung. Detected at
 CC lower levels in adult peripheral blood leukocytes, lung, and in
 CC fetal muscle, heart, kidney, small intestine and skin. Detected in
 CC T-cells, B-cells and monocytes. In T-cells expression is highest
 CC in Th0 cells, intermediate in Th2 cells and lower in Th1 cells.
 CC -1- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.
 CC -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-25 IS THE INITIATOR.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC -----
 DR EMBL: AF322069; AAG38115.1; -;
 DR EMBL: AY043489; AAK74193.1; -;
 DR EMBL: BC016420; AAH16420.1; -;
 DR HSSP: 014763; 1D0G.
 DR MGD: MGI:2151075; Trnfzf21.
 DR InterPro: IPR000488; Death.
 DR InterPro: IPR001368; TNFR_c6.
 DR Pfam: PF00020; TNFR_c6; 4.
 DR ProDom: PD000771; TNFR_c6; 1.
 DR SMART: SM00005; DEATH; 1.
 DR SMART: SM00208; TNFR; 4.
 DR PROSITE: PS50017; DEATH DOMAIN; 1.
 DR PROSITE: PS50052; TNFR_NGFR_1; 1.
 DR PROSITE: PS50050; TNFR_NGFR_2; 1.
 KW Receptor; Apoptosis; Transmembrane; Glycoprotein; Repeat; Signal.
 FT SIGNAL 1 41
 FT CHAIN 42 655
 FT DOMAIN 42 655
 FT TRANSMEM 350 349
 FT DOMAIN 371 370
 FT DOMAIN 415 498
 FT REPEAT 50 88
 FT TNFR-CYS 1.

FT REPEAT 90 131
 FT REPEAT 133 167
 FT REPEAT 170 211
 FT DISULFID 67 80
 FT DISULFID 70 88
 FT DISULFID 91 106
 FT DISULFID 109 123
 FT DISULFID 113 131
 FT DISULFID 133 144
 FT DISULFID 150 168
 FT DISULFID 171 186
 FT DISULFID 192 211
 FT CARBOHYD 82 82
 FT CARBOHYD 141 141
 FT CARBOHYD 252 252
 FT CARBOHYD 257 257
 FT CARBOHYD 278 278
 FT CARBOHYD 289 289
 FT CONFLICT 352 352
 FT CONFLICT 523 523
 SQ SEQUENCE 655 AA; 71982 MW; 5EC7C51C7C99EFF7 CRC64;
 Query Match 10.8%; Score 243; DB 1; Length 655;
 Best Local Similarity 26.5%; Pred. No. 2; 9e-07;
 Matches 76; Conservative 37; Mismatches 124; Indels 50; Gaps 13;
 QY 57 ICSSRCPGGTYYSAKSCSRIDYVATCAENSNEHNNYLTICOLC-RPCD-PWGLGEETA 114
 DB 65 LNCDCPKPACTYVSEHCTNMSLRVCSSCPAGTFREHNGIERHCDCQPCWPWMI---ERL 121
 QY 115 PCTSKRKTGCRQROPGME-----CAAWALCTHCELLSDCPGTEALDKDEVGKGN--HC 167
 DB 122 PCAALTRDREICPPGMQNSGTCAPHTV-----CPVGNGVKK---GTENDVVC 168
 QY 168 VPKAGHFQNTSSPSARCPHRCENQGLVEAPGTAQSDTCKNPL--EPLPEWSSGL 225
 DB 169 KCCARGTFFDVSYSVKCAHFTCLGQNLLEVYKPKETDNCVMRLFFSSTNPSSGTV 228
 QY 226 LKRRPQG-EGPNPVAGSWMPRAHPTFPDLVQPLPLISGDVSYSGTLAAPPLGNGVQ 284
 DB 229 TFSHPHMHSHVDPSSSTYEPQGNNS-----TDSNSTSVKRVKPSG-IEETVPD 277
 QY 285 QOSPLD-----LTREPLEPQEQSQAAGHTNGITHVGVGSMWTIG 323
 DB 278 NNSSTGKRGKTRNTPNPQV---THQAPHHHILKLPSSMEARG 321
 RESULT 14
 TR14_HUMAN
 ID TR14_HUMAN STANDARD; PRT; 283 AA.
 AC Q92956; Q9UM65; Q96J31; Q8WKRI;
 DT 16-OCT-2001 (Rel. 40; Created)
 DT 16-OCT-2001 (Rel. 40; Last sequence update)
 DT 15-JUN-2002 (Rel. 41; Last annotation update)
 DE Tumor necrosis factor receptor superfamily member 14 precursor
 DE (Herpesvirus entry mediator A) (Tumor necrosis factor receptor-like 2)
 DE (TR2).
 GN TNFRSF14 OR HVEM OR HVEA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cervical adenocarcinoma;
 RX MEDLINE=97053782; PubMed=8898196;
 RA Montgomery R.L., Warner M.S., Lum B.J., Spear P.G.;
 RT "Herpes simplex virus-1 entry into cells mediated by a novel member of
 RT the TNF/NGF receptor family.";
 RL Cell 87:427-436(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97306336; PubMed=9162061;

RA Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,
 RA Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,
 RA Suggs S., Boyle W.J.:
 RT "Osteoprotegerin: a novel secreted protein involved in the regulation
 RT of bone density.";
 RL Cell 89:309-319(1997).
 [2]
 RP SEQUENCE FROM N.A., AND VARIANTS PRO-138; ARG-161; ASP-165; ALA-288
 RP AND ARG-296.
 RC STRAIN-129/Ola, and NIH Swiss: TISSUE=Fibroblast;
 RA MEDLINE=96382527; Pubmed=9714833;
 RA Mizuno A., Murakami A., Nakagawa N., Yasuda H., Tsuda E., Morinaga T.,
 RA Higashio K.:
 RT "Structure of the mouse osteoclastogenesis inhibitory factor (OCIF)
 RT gene and its expression in embryogenesis.";
 RL Gene 215:339-343(1998).
 RP FUNCTION.
 RX MEDLINE=21060987; Pubmed=10952716;
 RA Min H., Morony S., Sarosi I., Dunstan C.R., Capparelli C., Scully S.,
 RA Van G., Kaufman S., Kostenuik P.J., Lacey D.L., Boyle W.J.,
 RA Simonet W.S.:
 RT "Osteoprotegerin reverses osteoporosis by inhibiting endosteal
 RT osteoclasts and prevents vascular calcification by blocking a process
 RT resembling osteoclastogenesis.";
 RL J. Exp. Med. 192:463-474(2000).
 CC -I- FUNCTION: Acts as decoy receptor for RANKL and thereby neutralizes
 CC its function in osteoclastogenesis. Inhibits the activation of
 CC osteoclasts and promotes osteoclast apoptosis in vitro. Bone
 CC homeostasis seems to depend on the local RANKL/OPG ratio. May also
 CC play a role in preventing arterial calcification. May act as decoy
 CC receptor for TRAIL and protect against apoptosis. TRAIL binding
 CC blocks the inhibition of osteoclastogenesis.
 CC -I- SUBUNIT: Homodimer.
 CC -I- SUBCELLULAR LOCATION: Secreted.
 CC -I- TISSUE SPECIFICITY: Highly expressed in liver, lung, stomach,
 CC intestines and calvaria. Highly expressed in decidua and placenta,
 CC and in embryo.
 CC -I- DEVELOPMENTAL STAGE: Detected in embryo at high levels on day 7,
 CC whereas expression decreases at day 11 and increases from day 15
 CC to 17. On day 15 found in developing bone primordia,
 CC brachiocephalic artery and ductus arteriosus, left main bronchus,
 CC abdominal aorta and midgut.
 CC -I- INDUCTION: Upregulated by TGF-beta and estrogens. Downregulated by
 CC 1,25-dihydroxyvitamin D3 and parathyroid hormone.
 CC -I- SIMILARITY: CONTAINS 2 DEATH DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 DEATH DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U94331; AAB53708.1; -;
 DR EMBL: AB013898; BAA28269.1; -;
 DR EMBL: AB013903; BAA33388.1; -;
 DR EMBL: AB013899; BAA33388.1; JOINED.
 DR EMBL: AB013900; BAA33388.1; JOINED.
 DR EMBL: AB013901; BAA33388.1; JOINED.
 DR EMBL: AB013902; BAA33388.1; JOINED.
 DR HSSP: P25942; ICDF.
 DR MGP: MGI:109587; TNFRsf1b.
 DR InterPro: IPR00488; Death.
 DR InterPro: IPR00368; TNFR_c6.
 DR Pfam: PF00020; TNFR_c6; 3.
 DR ProDom: PD000771; TNFR_c6; 1.
 DR SMART: SM00208; TNFR; 4.
 DR SMART: SM00005; DEATH; 1.
 DR PROSITE: PS50017; DEATH_DOMAIN; 1.
 DR PROSITE: PS00652; TNFR_NGFR_1; 1.

DR PROSITE: PS50050; TNFR_NGFR_2; 2.
 KW Receptor; Apoptosis; Glycoprotein; Repeat; Signal; Polymorphism.
 FT SIGNAL 1 21
 FT CHAIN 22 401
 FT
 FT REPEAT 24 62
 FT REPEAT 65 105
 FT REPEAT 107 142
 FT REPEAT 145 185
 FT DOMAIN 198 269
 FT DOMAIN 283 365
 FT SITE 400 400
 FT DISULFID 41 54
 FT DISULFID 44 62
 FT DISULFID 65 80
 FT DISULFID 83 97
 FT DISULFID 87 105
 FT DISULFID 107 118
 FT DISULFID 124 142
 FT DISULFID 145 160
 FT DISULFID 166 185
 FT CARBOHYD 98 98
 FT CARBOHYD 165 165
 FT CARBOHYD 178 178
 FT CARBOHYD 289 289
 FT VARIANT 138 138
 FT
 FT VARIANT 161 161
 FT
 FT VARIANT 165 165
 FT
 FT VARIANT 288 288
 FT
 FT VARIANT 296 296
 FT
 SO SEQUENCE 401 AA; 45923 MM; CAA6102D3B312470 CRC64;
 Query Match 10.9%; Score 244; DB 1; Length 401;
 Best Local Similarity 26.9%; Pred. No. 1.7e-07;
 Matches 54; Conservative 32; Mismatches 93; Indels 22; Gaps 6;
 Oy 18 VLGLFGLLAASQPOAVPYASENOTCRDQEKYEYEPQ--HRICSCRCPGTYSAKGSR 75
 Db 9 LVLVDITIEWTQETPLP-----KYLHYDPENGHQLCKCAAGFYLKQHCYVR 57
 Oy 76 RPTVATCAENSYNENHWNTLTICQLCRP-CDEVMGLLEIAPCTSKRKTCRCQPGMFCFA 134
 Db 58 RRTLCVPCPDHSDYTSWMTSDCYVCSFYCKELQSVKO--ECNRTNRYCECEBGRY--- 112
 Oy 135 WALECTHCELLSDCPPTGFAELCKDEVGKGNHCVCKAGHPONTSPSARCOPHRCENQ 194
 Db 113 --LEIEFLCKHNSCRPGS-GVYQAGTPENRYCKKCPDGFSGERTSKRAPIKHTNCSFT 169
 Oy 195 GLVEAAPGTASDTTCKNPLE 215
 Db 170 GLLLIQKGNATHDNCGNRE 190
 RESULT 13
 TR21_MOUSE STANDARD: PRT; 655 AA.
 AC Q9EPUS; Q91XH9; Q91W77;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tumor necrosis factor receptor superfamily member 21 precursor (TNFR-
 DE related death receptor-6) (Death receptor 6).
 GN TNFRSF21 OR DR6.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]

Query Match	Best Local Similarity	10.9%	Score 244.5	DB 1	Length 461
Matches 105; Conservative	25.2%	Pred. No. 1.8e-07	42; Mismatches 152; Indels 117; Gaps 22;		
SO SEQUENCE	461 AA;	50696 MM;	CD72361EC60C9DA3 CRC64;		
FT DOMAIN	362	447	DEATH.		
FT DISULFID	44	58	BY SIMILARITY.		
FT DISULFID	59	72	BY SIMILARITY.		
FT DISULFID	62	81	BY SIMILARITY.		
FT DISULFID	84	99	BY SIMILARITY.		
FT DISULFID	102	117	BY SIMILARITY.		
FT DISULFID	105	125	BY SIMILARITY.		
FT DISULFID	127	143	BY SIMILARITY.		
FT DISULFID	146	158	BY SIMILARITY.		
FT DISULFID	149	166	BY SIMILARITY.		
FT DISULFID	168	179	BY SIMILARITY.		
FT DISULFID	182	194	BY SIMILARITY.		
FT DISULFID	185	190	BY SIMILARITY.		
FT CARBOHYD	54	54	N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT CARBOHYD	86	86	N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT CARBOHYD	145	145	N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT CARBOHYD	151	151	N-LINKED (GLCNAC. . .) (POTENTIAL).		
Query Match	Best Local Similarity	10.9%	Score 244.5	DB 1	Length 461
Matches 105; Conservative	25.2%	Pred. No. 1.8e-07	42; Mismatches 152; Indels 117; Gaps 22;		
SO SEQUENCE	461 AA;	50696 MM;	CD72361EC60C9DA3 CRC64;		
FT DOMAIN	362	447	DEATH.		
FT DISULFID	44	58	BY SIMILARITY.		
FT DISULFID	59	72	BY SIMILARITY.		
FT DISULFID	62	81	BY SIMILARITY.		
FT DISULFID	84	99	BY SIMILARITY.		
FT DISULFID	102	117	BY SIMILARITY.		
FT DISULFID	105	125	BY SIMILARITY.		
FT DISULFID	127	143	BY SIMILARITY.		
FT DISULFID	146	158	BY SIMILARITY.		
FT DISULFID	149	166	BY SIMILARITY.		
FT DISULFID	168	179	BY SIMILARITY.		
FT DISULFID	182	194	BY SIMILARITY.		
FT DISULFID	185	190	BY SIMILARITY.		
FT CARBOHYD	54	54	N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT CARBOHYD	86	86	N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT CARBOHYD	145	145	N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT CARBOHYD	151	151	N-LINKED (GLCNAC. . .) (POTENTIAL).		
Query Match	Best Local Similarity	10.9%	Score 244.5	DB 1	Length 461
Matches 105; Conservative	25.2%	Pred. No. 1.8e-07	42; Mismatches 152; Indels 117; Gaps 22;		
SO SEQUENCE	461 AA;	50696 MM;	CD72361EC60C9DA3 CRC64;		
FT DOMAIN	362	447	DEATH.		
FT DISULFID	44	58	BY SIMILARITY.		
FT DISULFID	59	72	BY SIMILARITY.		
FT DISULFID	62	81	BY SIMILARITY.		
FT DISULFID	84	99	BY SIMILARITY.		
FT DISULFID	102	117	BY SIMILARITY.		
FT DISULFID	105	125	BY SIMILARITY.		
FT DISULFID	127	143	BY SIMILARITY.		
FT DISULFID	146	158	BY SIMILARITY.		
FT DISULFID	149	166	BY SIMILARITY.		
FT DISULFID	168	179	BY SIMILARITY.		
FT DISULFID	182	194	BY SIMILARITY.		
FT DISULFID	185	190	BY SIMILARITY.		
FT CARBOHYD	54	54	N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT CARBOHYD	86	86	N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT CARBOHYD	145	145	N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT CARBOHYD	151	151	N-LINKED (GLCNAC. . .) (POTENTIAL).		
Query Match	Best Local Similarity	10.9%	Score 244.5	DB 1	Length 461
Matches 105; Conservative	25.2%	Pred. No. 1.8e-07	42; Mismatches 152; Indels 117; Gaps 22;		
SO SEQUENCE	461 AA;	50696 MM;	CD72361EC60C9DA3 CRC64;		
FT DOMAIN	362	447	DEATH.		
FT DISULFID	44	58	BY SIMILARITY.		
FT DISULFID	59	72	BY SIMILARITY.		
FT DISULFID	62	81	BY SIMILARITY.		
FT DISULFID	84	99	BY SIMILARITY.		
FT DISULFID	102	117	BY SIMILARITY.		
FT DISULFID	105	125	BY SIMILARITY.		
FT DISULFID	127	143	BY SIMILARITY.		
FT DISULFID	146	158	BY SIMILARITY.		
FT DISULFID	149	166	BY SIMILARITY.		
FT DISULFID	168	179	BY SIMILARITY.		
FT DISULFID	182	194	BY SIMILARITY.		
FT DISULFID	185	190	BY SIMILARITY.		
FT CARBO					

KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal;
 KW Alternative splicing.
 FT SIGNAL 1 19
 FT CHAIN 20 289
 FT DOMAIN 20 193
 FT TRANSMEM 194 215
 FT DOMAIN 216 289
 FT REPEAT 25 60
 FT REPEAT 61 103
 FT REPEAT 104 144
 FT REPEAT 145 187
 FT DISULFID 26 37
 FT DISULFID 38 51
 FT DISULFID 41 59
 FT DISULFID 62 77
 FT DISULFID 83 103
 FT DISULFID 105 119
 FT DISULFID 111 116
 FT DISULFID 125 143
 FT CARBOHYD 153 153
 FT VARSPLIC 166 203
 FT VARSPLIC 204 289
 FT VARSPLIC 187 216
 FT VARSPLIC 216 234
 FT VARSPLIC 235 289
 FT VARSPLIC 216 222
 FT VARSPLIC 223 289
 SQ SEQUENCE 289 AA; 32111 MW; C791CB6D2FEA574E CRC64;
 Query Match 11.3%; Score 253; DB 1; Length 289;
 Best Local Similarity 25.5%; Pred. No. 3.8e-08;
 Matches 82; Conservative 38; Mismatches 128; Indels 74; Gaps 13;
 QY 8 SAPGLA--WGPLVLGLFGLLAQPOAVPYASENQTRDQDEKEYEYEPQHRICSRCPG 65
 DB 3 SLPRICALMGCL-----LTAVHLGQCV-----TCSDD-KQYLHGOQ--CDLQCPG 44
 QY 66 TYVSAKCSRIRDTVCATCAENSYNEHMYLITCOLCRPDPMVLEETAPCTSKRKTQCR 125
 DB 45 SRTLSHCALERTQCHPDSCGFSAQNNREIFCHQHRCEPQGLRVKKEGTAESDVTCT 104
 QY 126 CQPGKCAAMALECTHCELLSPCPG-----TEALKDEYVGKGNHCYPCRAHQNNTSS 180
 DB 105 KEGQHCT--SKDCACAOHPPCIPGFGVEMEMATEETTVV-----CHPCPVGFFSNQSS 156
 QY 181 PSARCOPTRCENOGILVAAPGTAOSDTTC--KNLEPLV--PREMGSILK----- 227
 DB 157 LFEKCYPTSCEDKNLEVLQKGTSTQNTVICGLKSRMLLVIPVWGLITITFGVFLYIK 216
 QY 228 ---RRPGGEGPNVAGSWEPKRAHYPFDLVQPLPLISGDVSPVSTGLPAAPVLEAGVQ 284
 DB 217 KVKKRPKNENMLPBARQDDPEMEDYPG-----INTAPAVQDTLHGQCPV 262
 QY 285 QOSPLDLTRPQLPEGEOOVA 306
 DB 263 TQ-----EDGKESRIS 273
 RESULT 10
 TNRS_BOVIN STANDARD: PRT; 269 AA.
 AC Q28203;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tumor necrosis factor receptor superfamily member 5 precursor
 DE (CD40L receptor) (B-cell surface antigen CD40) (Fragment).
 GN TNRSFS OR CD40.

OS Bos taurus (Bovine).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC Bovidae; Bovinae; Bos.
 CC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97281252; PubMed=9135560;
 RA Hiran A., Brown W.C., Estes D.M.;
 RT "Cloning, expression and biological function of the bovine CD40
 RT homologue: role in B-lymphocyte growth and differentiation in
 RT cattle."
 RL Immunology 90:294-300(1997).
 CC -1- FUNCTION: Receptor for TNRSF5/CD40L.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U57745; AAC48710.1; -.
 CC HSSP: P25942; ICDF.
 DR InterPro: IPR001368; TNFR_C6.
 DR Pfam: PF000020; TNFR_C6; 4.
 DR ProDom: PD000771; TNFR_C6; 1.
 DR SMART: SM00652; TNFR_NGFR_1; 1.
 DR PROSITE: PS00652; TNFR_NGFR_1; 1.
 DR PROSITE: PS00650; TNFR_NGFR_2; 1.
 KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 269
 FT DOMAIN 20 193
 FT TRANSMEM 194 215
 FT DOMAIN 216 269
 FT REPEAT 25 60
 FT REPEAT 61 103
 FT REPEAT 104 144
 FT REPEAT 145 187
 FT DISULFID 26 37
 FT DISULFID 38 51
 FT DISULFID 41 59
 FT DISULFID 62 77
 FT DISULFID 83 103
 FT DISULFID 105 119
 FT DISULFID 111 116
 FT DISULFID 125 143
 FT CARBOHYD 153 153
 FT CARBOHYD 180 180
 FT NON_TER 269
 SQ SEQUENCE 269 AA; 29983 MW; 746903F30F95F387 CRC64;
 Query Match 11.2%; Score 251.5; DB 1; Length 269;
 Best Local Similarity 33.1%; Pred. No. 4.3e-08;
 Matches 55; Conservative 21; Mismatches 81; Indels 9; Gaps 4;
 QY 47 EKEYEPQHRICSRCPGTVSAKSRIRDTVCATCAENSYNEHMYLITCOLCRPCDP 106
 DB 28 EKQY--PVNSLCCDLCPGQKLVNCTEVSSTEQSCGKGFELTWNKXCKCHERYCNP 85
 QY 107 VMLGELIAPCTSKRKTQCRCPGFMCAAMALECTHCELLSDCPGTEALKDEYVGK--N 164
 DB 86 NLGLRIQSEGTLNDTITICVGEQCHTSHT--CESCTPHSLCLPQFGVK--QIATGLLD 140
 QY 165 NHCVCKRAGHQNNTSSPARCOPTRCENOGILVAAPGTAOSDTTC 210
 DB 141 TVCEPCPLGFGVSNVSAPEKCHRWTSCKRGLVDEHVGNTKTDVVC 186

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AF018253; AAB86809.1; -
 DR HSSP; P25942; ICDF.
 DR Genew; HGNC:11908; TNFRSF11A.
 DR MIM; 603499; -
 DR MIM; 174810; -
 DR MIM; 602080; -
 DR InterPro; IPR001368; TNFR_c6.
 DR Pfam; PF00020; TNFR_c6; 4.
 DR ProDom; PD000771; TNFR_c6; 1.
 DR SMART; SM00208; TNFR; 4.
 DR PROSITE; PS00652; TNFR_NGFR_1; 1.
 DR PROSITE; PS00505; TNFR_NGFR_2; 1.
 KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal; Polymorphism;
 Disease mutation.
 FT SIGNAL 1 29
 FT CHAIN 30 616
 FT DOMAIN 30 212 POTENTIAL.
 FT TRANSMEM 213 233 TUMOR NECROSIS FACTOR RECEPTOR
 FT DOMAIN 234 616 SUPERFAMILY MEMBER 11A.
 FT REPEAT 34 68 EXTRACELLULAR (POTENTIAL).
 FT REPEAT 71 112 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 114 151 TNFR-CYS 1.
 FT REPEAT 154 194 TNFR-CYS 2.
 FT DISULFID 34 46 TNFR-CYS 3.
 FT DISULFID 47 60 TNFR-CYS 4.
 FT DISULFID 50 68 BY SIMILARITY.
 FT DISULFID 71 86 BY SIMILARITY.
 FT DISULFID 92 112 BY SIMILARITY.
 FT DISULFID 114 127 BY SIMILARITY.
 FT DISULFID 133 151 BY SIMILARITY.
 FT CARBOHYD 105 105 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 174 174 N-LINKED (GLCNAC...) (POTENTIAL).
 FT VARIANT 21 21 L -> LALLCAL (IN PDB2).
 FT VARIANT 21 21 /FTID-VAR_011516.
 FT VARIANT 21 21 L -> LALLCAL (IN FEO).
 FT VARIANT 21 21 /FTID-VAR_011517.
 FT VARIANT 192 192 A -> V.
 FT VARIANT 192 192 /FTID-VAR_011518.
 FT SEQUENCE 616 AA; 66033 MW; E3DE9A7A08196F81 CRC64;
 SQ
 Query Match 13.3%; Score 298; DB 1; Length 616;
 Best Local Similarity 31.1%; Pred. No. 2.1e-10;
 Matches 76; Conservative 31; Mismatches 97; Indels 40; Gaps 7;

ID TNRS_MOUSE STANDARD; PRT; 289 AA.
 AC P27512; Q99NE0; Q99NE1; Q99NE2; Q99NE3;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tumor necrosis factor receptor superfamily member 5 precursor
 DE (CD40L receptor) (B-cell surface antigen CD40) (BP50) (CDw40).
 GN TNFRSF5 OR CD40.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP MEDLINE=92105763; PubMed=1370315;
 RA Torres R.M., Clark E.A.,
 RT "Differential increase of an alternatively polyadenylated mRNA
 species of murine CD40 upon B lymphocyte activation.";
 RL J. Immunol. 148:620-626(1992).
 RN [2]
 RP REVISIONS.
 RC STRAIN-BALB/c;
 RA Torres R.M.;
 RL Submitted (SEP-1996) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM I).
 RC STRAIN-BALB/c; TISSUE=Liver;
 RA MEDLINE=93094586; PubMed=1281194;
 RX Gimaldi J.C., Torres R., Kozak C.A., Chang R., Clark E.A.,
 RA Howard M., Cockayne D.A.;
 RT "Genomic structure and chromosomal mapping of the murine CD40 gene.";
 RL J. Immunol. 149:3921-3926(1992).
 RN [4]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING (ISOFORMS II; III; IV AND V).
 RA MEDLINE=21117110; PubMed=1172023;
 RX Tone M., Tone Y., Fairchild P.J., Wykes M., Waldmann H.;
 RT "Regulation of CD40 function by its isoforms generated through
 alternative splicing.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:1751-1756(2001).
 CC - FUNCTION: Receptor for TNFSF5/CD40L.
 CC - SUBCELLULAR LOCATION: Type I membrane protein (isoforms I, III, IV and V); Secreted (isoform II).
 CC - ALTERNATIVE PRODUCTS: 5 isoforms; I (shown here), II, III, IV and V; are produced by alternative splicing.
 CC - SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

RA Nakagawa N., Kinoshita M., Yamaguchi K., Shima N., Yasuda H., Yano K.,
 RA Morinaga T., Higashio K.;
 RT "RANK is the essential signaling receptor for osteoclast
 differentiation factor in osteoclastogenesis.";
 RL Biochem. Biophys. Res. Commun. 253:395-400(1998).
 CC -1- FUNCTION: Receptor for TNFSF11/RANKL/TRANCE/OPGL; essential for
 CC RANKL-mediated osteoclastogenesis. Involved in the regulation of
 CC interactions between T-cells and dendritic cells.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
 CC -1- TISSUE SPECIFICITY: UBIQUITOUS EXPRESSION WITH HIGH LEVELS IN
 CC TRABECULAR BONE, THYMUS, SMALL INTESTINE, LUNG, BRAIN AND KIDNEY.
 CC WEAKLY EXPRESSED IN SPLEEN AND BONE MARROW.
 CC -1- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF019046; AAB96810.1; -
 DR EMBL: BC019185; AAH19185.1; -
 DR HSSP: P25942; ICDP.
 DR MGD: MGI:1314891; Trnrsfl1a.
 DR InterPro: IPR001368; TNFR_C6.
 DR Pfam: PF000020; TNFR_C6; 3.
 DR Prodom: P0000771; TNFR_C6; 1.
 DR SMART: SM00208; TNFR_4.
 DR PROSITE: PS00652; TNFR_NGFR_1; 1.
 DR PROSITE: PS50050; TNFR_NGFR_2; 1.
 KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal.
 FT SIGNAL 1 30
 FT CHAIN 1 625 TUMOR NECROSIS FACTOR RECEPTOR
 FT DOMAIN 31 214 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 215 234 POTENTIAL.
 FT DOMAIN 235 625 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 35 69 TNFR-CYS 1.
 FT REPEAT 72 113 TNFR-CYS 2.
 FT REPEAT 115 152 TNFR-CYS 3.
 FT REPEAT 155 195 TNFR-CYS 4.
 FT DISULFID 35 47 BY SIMILARITY.
 FT DISULFID 48 61 BY SIMILARITY.
 FT DISULFID 51 69 BY SIMILARITY.
 FT DISULFID 72 87 BY SIMILARITY.
 FT DISULFID 93 113 BY SIMILARITY.
 FT DISULFID 115 128 BY SIMILARITY.
 FT DISULFID 134 152 BY SIMILARITY.
 FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 494 494 R -> K (IN REF. 2).
 SQ SEQUENCE 625 AA: 66621 MW: F8C1872E9511D8E CRK64;
 Query Match 13.9%; Score 312; DB 1; Length 625;
 Best Local Similarity 32.0%; Pred. No. 3.3e-11;
 Matches 79; Conservative 27; Mismatches 101; Indels 40; Gaps 7;
 Oy 18 VLGLFELLASGP--QAVPYASENTCRDQEKYEYEPHRCICSCRPCTVSACSAI 75
 Db 15 LIALCVLLPDLVTLQVTPCTOE-----RHYEHGGR-CCSRCEGKYLSSCTPT 64
 Oy 76 RPTVCATCAENSYNEHWNLTTCQLCRPCDPVWGLEIEIAPCTSKRTQCRCPGMCAAM 135
 Db 65 SOSVCLCPGDEYLDTWNEDKCLLHKVDAGKALVAVDPGNHTARRCACIAGYH--W 121
 Oy 136 ALECHGCELLSDCPGETEKLDEYGNKNNHCVPCKAGHFONTSPSARCOPTRCENOG 195
 Db 122 NSSCCECCRRNTCEAPGFGQHPLQLNK-DTVCCTPLGFFSDFVSTSDCKKPTNCTLLG 180
 Oy 196 IVEAARGTASDPTCKNPLPLEPPEMSSGLKRRRQSGEPNPVAGSMERPKAHPPYPDV 255

Db 181 KLEAHQGTESDPVC-----SSSMTLRRP-----PKNAAYLPGLI 216
 Oy 256 OPLPLPS 262
 Db 217 VLLPLPS 223
 RESULT 8
 ID TR1_HUMAN STANDARD; PRT; 616 AA.
 AC Q9Y6Q6;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Tumor necrosis factor receptor superfamily member 11A precursor
 DE (Receptor activator of NF-kB) (Osteoclast differentiation factor
 DE receptor) (ODFR).
 GN TNFRSF11A OR RANK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Dendritic cell;
 RX MEDLINE=98032977; PubMed=9367155;
 RA Anderson D.M., Maraskovsky E., Billingsley W.L., Dougall W.C.,
 RA Tomeles M.E., Roux E.R., Teepe M.C., Dubose R.F., Cosman D.,
 RA Galibert L.;
 RT "A homologue of the TNF receptor and its ligand enhance T-cell growth
 RT and dendritic-cell function.";
 RL Nature 390:175-179(1997).
 RN [2]
 RP MEDLINE=99097247; PubMed=9878548;
 RX Nakagawa N., Kinoshita M., Yamaguchi K., Shima N., Yasuda H., Yano K.,
 RA Morinaga T., Higashio K.;
 RT "RANK is the essential signaling receptor for osteoclast
 RT differentiation factor in osteoclastogenesis.";
 RL Biochem. Biophys. Res. Commun. 253:395-400(1998).
 RN [3]
 RP VARIANT PEO 16-L--L-21 DUPL, VARIANT PDB2 13-A--L-21 DUPL, AND VARIANT
 RP V-192.
 RX MEDLINE=20082806; PubMed=10615125;
 RA Hughes A.E., Ralston S.H., Marken J., Bell C., Macpherson H.,
 RA Wallace R.G.H., van Hul W., Whyte M.P., Nakatsuka K., Hoy L.,
 RA Anderson D.M.;
 RT "Mutations in TNFRSF11A, affecting the signal peptide of RANK, cause
 RT familial expansile osteolysis.";
 RL Nat. Genet. 24:45-48(2000).
 CC -1- FUNCTION: Receptor for TNFSF11/RANKL/TRANCE/OPGL; essential for
 CC RANKL-mediated osteoclastogenesis. Involved in the regulation of
 CC interactions between T-cells and dendritic cells.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
 CC -1- TISSUE SPECIFICITY: UBIQUITOUS EXPRESSION WITH HIGH LEVELS IN
 CC SKELETAL MUSCLE, THYMUS, LIVER, COLON, SMALL INTESTINE AND ADRENAL
 CC GLAND.
 CC -1- DISEASE: DEFECTS IN TNFRSF11A ARE THE CAUSE OF FAMILIAL EXPANSILE
 CC OSTEOLYSIS (PEO), A RARE AUTOSOMAL DOMINANT BONE DISORDER
 CC CHARACTERIZED BY FOCAL AREAS OF INCREASED BONE REMODELING. THE
 CC OSTEOCLYTIC LESIONS DEVELOP USUALLY IN THE LONG BONES DURING EARLY
 CC ADULTHOOD. PEO IS OFTEN ASSOCIATED WITH EARLY ONSET DEAFNESS AND
 CC LOSS OF DENTITION.
 CC -1- DISEASE: DEFECTS IN TNFRSF11A ARE A CAUSE OF FAMILIAL PAGET
 CC DISEASE OF BONE, ALSO KNOWN AS PAGET DISEASE OF BONE 2 (PDB2). IT
 CC IS A BONE REMODELING DISORDER WITH CLINICAL SIMILARITIES TO PEO.
 CC UNLIKE PEO, HOWEVER, AFFECTED INDIVIDUALS HAVE INVOLVEMENT OF THE
 CC AXIAL SKELETON WITH LESIONS IN THE SPINE, PELVIS AND SKULL.
 CC -1- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

	FT	DISULFID	125	143	N-LINKED (GLCNAC. . .) (POTENTIAL).
	FT	CARBOHYD	153	153	N-LINKED (GLCNAC. . .) (POTENTIAL).
	FT	CARBOHYD	180	180	SCENRDLVVOAGNCKTDYVCQGDRLRALVALPIIFG ->
	FT	VARSPLIC	166	203	RPSASAPSGDDPHNLDPVCHPLIGALTYOKGGOEANO
	FT	VARSPIC			(IN ISOFORM II)
	FT	VARSPIC			(IN ISOFORM II)
	FT	VARSPIC			(IN ISOFORM II).
	FT	VARSPIC			C-> R (IN H10M3).
	FT	VARSPIC			/FtId=VAR_013628.
	SO	SEQUENCE	277 AA:	30619 MW:	BC8776EC2CA5680 CRC64;
		Query Match	14.0%;	Score 314.5;	DB 1;
		Best Local Similarity	29.8%;	Pred. No. 1.2e-11;	
		Matches 92;	Conservative	38;	Mismatches 120; Indels 59; Gaps
OY		16 PVLGLFG-FGLAASOPAAVPYASSENQGRDOEKYEYEPQHRICCSRCPPTGYSAKSR	74		
Db		5 PLQCVLMGCLLTAVHRE--PFTA-----CR-EKOYLINSQ--CCSLQPOQKLVSDCTE	53		
OY		75 IRTVTATCAENSYNENHNWYLIITCOLCRCDDVMGLEIAPTSCRKTQRCROCPGFCAA	134		
Db		54 FTETELPCGESEFLDTWRETHGHQHKYCDDPLGNVQOKTSETDITCTCEEBNHCTS	113		
OY		135 WALECFHCCELLISCPPTGETEALKDEYKGKNNNHCPCKACHPONTSSPSARCOPHTCENQ	194		
Db		114 EA-CESCVLHRCSGFGVGK-QIANGVDITCEPCPVGFNSVAFEKCHPMYSCTEK	170		
OY		195 GLVEAAPGTASDPTCKNPLELP-----PENSGSL-----LKRPDGEENP	237		
Db		171 DLVVOAGTGTKTTPYVC-GPDRLRALVPIPIFGILFAILLVLFIKAKVPTPKAHHP	229		
OY		238 VAGSWEPKAPHFPDLYQPLIPISGDVSVSTGTLPAAPVLEAGVPOOQSPLDLTREPOL	297		
Db		230 ---KQEQEIN--FPD-----DLPSNTAAAPVQETLHGCGVTO-----	263		
OY		298 EPGEOSOVA	306		
Db		264 EDGKEKRIS	272		
<hr/>					
RESULT 7					
ID	TRIM_MOUSE	STANDARD;	PRT;	625 AA.	
AC	035305; Q8VCT7;				
DT	16-OCT-2001 (Rel. 40, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Tumor necrosis factor receptor superfamily member 11A precursor				
DE	(Receptor activator of NF-kB) (osteoclast differentiation factor				
DE	receptor) (ODFR).				
GN	TNFRSF1A OR RANK.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_Taxid=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Fetal liver;				
RX	MEDLINE=98032977; PubMed=93671155;				
RA	Anderson D.M., Maraskovsky E., Billingsley W.L., Dougall W.C.,				
RA	Tometz M.E., Roux E.R., Teepe M.C., Dubose R.F., Cosman D.,				
RA	Gallibert L.,				
RT	"A homologue of the TNF receptor and its ligand enhance T-cell growth				
RT	and dendritic-cell function."				
RL	Nature 390:175-179(1997).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Mammary gland;				
RA	Straussberg R.,				
RL	Submitted (Dec-2001) to the EMBL/GenBank/DDBJ databases.				
RN	[3]				
RP	FUNCTION.				
RX	MEDLINE=99097247; PubMed=9878548;				

CC from colon, stomach, rectum, esophagus and in SW480 colon
CC carcinoma cells.

CC -1- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL: AF104419; AAD03056.1; -
DR EMBL: AF134240; AAD29688.1; -
DR EMBL: AF217796; AAF35244.1; -
DR EMBL: AF217793; AAF33685.1; -
DR EMBL: AF217794; AAF33686.1; -
DR EMBL: AL121845; CAC03668.1; -
DR EMBL: BC017065; AAH17065.1; -
DR Genew: HGNC:11921; TNFRSF6B.
DR MIM: 603361; -
DR HSSP: 014763; 1D0G.
DR Interpro: IPR001368; TNFR_C6.
DR Pfam: PF00020; TNFR_C6; 4.
DR ProDom: PD000771; TNFR_C6; 1.
DR SMART: SM00208; TNFR; 3.
DR PROSITE: PS00652; TNFR_NGFR_1; 2.
DR PROSITE: PS50050; TNFR_NGFR_2; 2.
DR Receptor: Apoptosis; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 29
FT CHAIN 30 300
FT REPEAT 31 70 TUMOR NECROSIS FACTOR RECEPTOR
FT REPEAT 72 113 SUPERFAMILY MEMBER 6B.
FT REPEAT 115 150 TNFR-CYS 1.
FT REPEAT 152 193 TNFR-CYS 3.
FT REPEAT 195 237 TNFR-CYS 4.
FT DISULFID 49 62 BY SIMILARITY.
FT DISULFID 52 70 BY SIMILARITY.
FT DISULFID 73 88 BY SIMILARITY.
FT DISULFID 91 105 BY SIMILARITY.
FT DISULFID 95 113 BY SIMILARITY.
FT DISULFID 115 126 BY SIMILARITY.
FT DISULFID 132 150 BY SIMILARITY.
FT DISULFID 153 168 BY SIMILARITY.
FT DISULFID 174 193 BY SIMILARITY.
FT CARBOHYD 173 173 N-LINKED (GLCNAc...) (POTENTIAL).
SQ SEQUENCE 300 AA: 32679 MW: F90AE33718449AF CRC64;

Query Match 14.2%; Score 319; DB 1; Length 300;
Best Local Similarity 33.0%; Pred. No. 6,9e-12;
Matches 88; Conservative 24; Mismatches 103; Indels 52; Gaps 12;

QY 6 ATSPAGLAPVILGFLGILAAQPOAV--PPYASENQCROEKEYEPOHRCISRC 62
DB 3 ALEGPILSLCLVLPALLPVPVAVGVAETPY-----PWRDA-----ETGERLVCAQC 52
QY 63 PPGTYVSACSRITRDVTCATCAENSYNEHMYLTITQLRCPDPVNG--LEETAPCTSKR 120
DB 53 PPGTFVQRCRDRSPTTCGPCPRHYTQFWNYL---ERCRCYVNLGGEDEEARACHAHV 109
QY 121 KTGRCRQPGMFC-AAMALCTHCELLSDCPGTEALKDEVGKGNHCYPCAKAGHQNTS 179
DB 110 NRACRCRTTFEFAHAGCLE--H---ASCPGA-GVIADGTPSONTCQCPGPGFSASS 162
QY 180 SPFARCPHTRCENGLVFAAPGTAOSDTCKN---PLEPLPP----- 219
DB 163 SSSGQCQPHRNCATGALANVRSSSHDTLCTSGTFPLSTRVGAEECEBRAVIDVARQ 222
QY 220 EMSGSLKR-----RPQEGEPNPAV 240
DB 223 DISIKRLQRLQALFAPEGMGPTPRAG 249

RESULT 6
ID TRN5_HUMAN
AC P25942; Q9BYU0;
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor receptor superfamily member 5 precursor
DE (CD40L receptor) (B-cell surface antigen CD40) (CDW40) (Bp50).
GN TNFRSF5 OR CD40.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP MEDLINE=89355608; PubMed=2475341;
RA Stamenkovic I., Clark E.A., Seed B.;
RT "A B-lymphocyte activation molecule related to the nerve growth
RT factor receptor and induced by cytokines in carcinomas.";
RL EMBO J. 8:1403-1410(1989).
[2]
RP SEQUENCE FROM N.A. (ISOFORM I).
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggaley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carter C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Ellington D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Levasialho M.H., Leverisha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McMay R., McMurray A.A.,
RA Milne S.A., Misty D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Philimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkhen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.A., Sycamore A.C., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
[3]
RP SEQUENCE FROM N.A. (ISOFORM II).
RX MEDLINE=21117110; PubMed=11172023;
RA Tone M., Tone Y., Fairchild P.J., Wykes M., Waldmann H.;
RT "Regulation of CD40 function by its isoforms generated through
RT alternative splicing.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:1751-1756(2001).
[4]
RP SEQUENCE FROM N.A. (ISOFORM I).
RC TISSUE=ovary;
RA Strausberg R.;
RL Submitted (Aug-2001) to the EMBL/Genbank/DBJ databases.
[5]
RP 3D-STRUCTURE MODELING OF 24-144.
RX MEDLINE=97189482; PubMed=9037712;
RA Bajorath J., Aruffo A.;
RT "Construction and analysis of a detailed three-dimensional model of
RT the ligand binding domain of the human B cell receptor CD40.";
RL Proteins 27:59-70(1997).
[6]
RP 3D-STRUCTURE MODELING OF 26-186 IN COMPLEX WITH CD40L.
RX MEDLINE=98266353; PubMed=9605317;

RA Jacob C.O., Liu J.;
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE OF 1-22 FROM N.A.
 RC TISSUE=Liver;
 RA Kissomerginis M., Fellows R., Feldmann M., Chernajovsky Y.;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Receptor with high affinity for TNFSF2/TNF-alpha and
 CC approximately 5-fold lower affinity for homotrimeric
 CC TNFSF1/Lymphotoxin-alpha (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M60469; AAA39752.1; -;
 DR EMBL: M59378; AAA40463.1; -;
 DR EMBL: U39488; AAA85021.1; -;
 DR EMBL: X87128; CAA60618.1; -;
 DR PIR: B38634; B38634.
 DR HSSP: P19438; INCF.
 DR MGD: MG11314883; Tnfisfld.
 DR InterPro: IPR01368; TNFR_c6.
 DR Pfam: PF00020; TNFR_c6; 4;
 DR PRODOM: PD000771; TNFR_c6; 1;
 DR SMART: SM00208; TNFR; 4;
 DR PROSITE: PS00652; TNFR_NGFR_1; 2;
 DR PROSITE: PS50050; TNFR_NGFR_2; 3;
 DR Receptor; Transmembrane; Glycoprotein; Repeat; Signal.
 KW SIGNAL
 FT 1 22
 FT CHAIN 23 474
 FT 23 258
 FT 23 258
 FT TRANSMEM 259 288
 FT DOMAIN 289 474
 FT REPEAT 399 77
 FT REPEAT 78 119
 FT REPEAT 120 164
 FT REPEAT 165 203
 FT DISULFID 40 54
 FT DISULFID 55 68
 FT DISULFID 58 76
 FT DISULFID 79 94
 FT DISULFID 97 111
 FT DISULFID 101 119
 FT DISULFID 121 127
 FT DISULFID 136 145
 FT DISULFID 139 163
 FT DISULFID 166 181
 FT CARBOHYD 69 69
 FT CARBOHYD 195 195
 FT SEQUENCE 474 AA; 50319 MW; 462EAE398CAD6363 CRC64;
 Query Match 14.6%; Score 327.5; DB 1; Length 474;
 Best Local Similarity 29.5%; Pred. No. 3.4e-12;
 Matches 101; Conservative 40; Mismatches 142; Indels 59; Gaps 16;

DB 114 QONRVACGEAGRYCALKTHSGSCGRCMLSKCGPGF-GVASSRAPNGNVLCRACAPGTF 172
 QY 177 NTSSPSARCPHRCENOGIVEAPGTAGSDTTCKNPLEPLPEMGSLKRRPQGEQPN 236
 DB 173 DTSSTFVCRPHRCS---ILAIQGNASTDVA-PESEPTLSAIRPTLYVSOPEPTRSQ 227
 QY 237 PVAGSWEPRAHRYFPLVPL-----LPTSGVSPVSTGLPAPVLEAG 281
 DB 228 PL--DDEPGSQT--PSILNLSSTPIIESTKGGLSLPLGLIVGYSLSL-----LMLG 278
 QY 282 -----VPOQSPDLDTREPOLE--PGESQVANGTNGIVH 314
 DB 279 LVNCTILVGRKKRPSCLGRDAKVPHPVDEKSDQAVGLEQGH 320
 RESULT 5
 TR6B_HUMAN STANDARD; PRT; 300 AA.
 AC 095407;
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tumor necrosis factor receptor superfamily member 6B precursor (Decoy
 GN receptor for Fas ligand) (Decoy receptor 3) (DCR3) (M68).
 OS Homo sapiens (human)
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal lung;
 RX MEDLINE=99087326; PubMed=9872321;
 RA Pitti R.M., Marsters S.A., Lawrence D.A., Roy M., Kischkel F.C.,
 RA Dowd P., Huang A., Donahue C.J., Sherwood S.W., Baldwin D.T.,
 RA Goddard A.D., Botstein D., Gurney A.L., Hillan K.J., Cohen R.L.,
 RA "Genomic amplification of a decoy receptor for Fas ligand in lung and
 RT colon cancer".
 RL Nature 396:699-703(1998).
 RN [2]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 30-35.
 RC TISSUE=Prostate;
 RX MEDLINE=99233915; PubMed=10318773;
 RA Yu K.-Y., Kwon B., Ni J., Zhai Y., Ebner R., Kwon B.S.;
 RA "A newly identified member of tumor necrosis factor receptor
 RT superfamily (TR6) suppresses liver mediated apoptosis".
 RL J. Biol. Chem. 274:13733-13736(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=20122600; PubMed=1065513;
 RA Bal C., Connolly B., Metzger M.L., Hilliard C.A., Liu X., Sandly V.,
 RA Soderman A., Gallows S.M., Liu Q., Austin C.P., Caskey C.T.;
 RA "Overexpression of M68/DCR3 in human gastrointestinal tract tumors
 RT independent of gene amplification and its location in a four-gene
 RT cluster".
 RL Proc. Natl. Acad. Sci. U.S.A. 97:1230-1235(2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Mathews L.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RA Straube R.;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Decoy receptor for the cytotoxic ligands TNFSF14/LIGHT
 CC and TNFSF6/FasL. Protects against apoptosis.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Detected in fetal lung, brain and liver.
 CC detected in adult stomach, spinal cord, lymph node, trachea,
 CC spleen, colon and lung. Highly expressed in several primary tumors


```

FT DOMAIN 245 415 CYTOPLASMIC (POTENTIAL).
FT REPEAT 42 81 TNFR-CYS 1.
FT REPEAT 82 124 TNFR-CYS 2.
FT REPEAT 125 170 TNFR-CYS 3.
FT REPEAT 171 213 TNFR-CYS 4.
FT DISULFID 43 58 BY SIMILARITY.
FT DISULFID 59 72 BY SIMILARITY.
FT DISULFID 62 80 BY SIMILARITY.
FT DISULFID 83 98 BY SIMILARITY.
FT DISULFID 101 116 BY SIMILARITY.
FT DISULFID 104 124 BY SIMILARITY.
FT DISULFID 126 132 BY SIMILARITY.
FT DISULFID 139 150 BY SIMILARITY.
FT DISULFID 142 169 BY SIMILARITY.
FT DISULFID 172 187 BY SIMILARITY.
FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 415 AA: 44956 MM: 298326566AEF61 CRC64;

```

```

Query Match 62.5%; Score 1404; DB 1; Length 415;
Best Local Similarity 63.9%; Pred. No. 1.8e-73;
Matches 265; Conservative 28; Mismatches 86; Indels 36; Gaps 4;

```

```

OY 1 MLPMATASAGLAMPVILGFLGLAASOPAVPVASENOTCRODEREYEPQHRICCS 60
DB 1 MRPLPRASSPCGLAMGPIILGLSLVASOPOLVPIRIENQTCDDDKETPEPMHIVCCS 60
OY 61 RCPGTVYSAKSRIRDTVCATCAENSYNEHNNVLTICLCRPDPVYGLAEIAPCTSKR 120
DB 61 RCPGEEVEFVAVCSRSQDTVCCTCPHNSYNEHNNVLTICLCRPDPVYGLAEIAPCTSKR 120
OY 121 KTCRCOPGMCAMALECHNC-ELLSDCPGFEALIKQEVGNGNHCVCARAGHONT 178
DB 121 KAEBCRCOPGMCVYLIDNECVCHCEERLVLCOPGTEAEVTDIEMDTGVNCVCRGHRONT 180
OY 179 SSPSARCOPTHRCENOGILVEAPGTAQSDFTCKNPLEP----- 216
DB 181 SSPARCOPTHRCIOGLVEAPGTSYSDITCKNPPRGAMLLAIIISLVLFLLFTTVL 240
OY 217 -----LPEPMS--GSLIKRRPGGEGPNVAGSWEPPKAMPYPDVQLPIGSDVSPV 268
DB 241 ACAMWRHPSLCRKLTGLTLKRHEGESESPCPA---PRADPHFDLLEPLPMGSDISPS 296
OY 269 STGLPAPVLEAGVPOQOSPLDLTRPOLERPEQSOVAGHNGIHVGGSGMTIGNIYIY 328
DB 297 PAGPPTASILEEVYLOQOSPLVQARELEAEPEGHQVAGHNGIHTVGGSTVIGNIYIY 356
OY 329 NGPVILGPPGPDLPATPEPPYPIPEEGDPPPGISLTPHQDGRAMHLAETEHCG 383
DB 357 NGPVILGTRGPDPPAPEPPYPIPEEGAPGPESELSTPYQDGRAMHLAETETLG 411

```

```

RESULT 3
TRIB_HUMAN STANDARD; PRT: 461 AA.
AC P20333; O16042;
ID 01-FEB-1991 (Rel. 17, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor receptor superfamily member 1B precursor (Tumor
DE necrosis factor receptor 2) (P80) (TNF-R2) (p75) (CD120b) (Etarcept)
DE [Contains: Tumor necrosis factor binding protein 2 (TNFRII)].
DE TNFRSF1B OR TNFR2 OR TNFBR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP MEDLINE FROM N.A.
RA MEDLINE-90260639; PubMed-2160731;
RA Smith C.A., Davis T., Anderson D., Solam L., Beckmann M.P., Jerzy R.,
RA Dower S.K., Cosman D., Goodwin R.G.;
RA "A receptor for tumor necrosis factor defines an unusual family of

```

```

RT cellular and viral proteins."
RL Science 248:1019-1023(1990).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE-91045991; PubMed-2172983;
RA Kohno T., Brewer M.T., Baker S.L., Schwartz P.E., King M.W.,
RA Hale K.K., Squires C.H., Thompson R.C., Vanice J.L.;
RT "A second tumor necrosis factor receptor gene product can shed a
RT naturally occurring tumor necrosis factor inhibitor."
RL Proc. Natl. Acad. Sci. U.S.A. 87:8331-8335(1990).
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE-96299745; PubMed-8661109;
RA Beltinger C.P., White P.S., Maris J.M., Sulman E.P., Jensen S.J.,
RA Lepaslier D., Stallard B.J., Goeddel D.V., Desautelle F.J.,
RA Brodeur G.M.;
RT "Physical mapping and genomic structure of the human TNFR2 gene."
RL Genomics 35:94-100(1996).
[4]
RP SEQUENCE OF 37-461 FROM N.A.
RX MEDLINE-91370690; PubMed-1966549;
RA Demic Z., Loetscher H., Gubler U., Pan Y.C., Lahm H.W., Gentz R.,
RA Brockhaus M., Lesslauer W.;
RT "Two human TNF receptors have similar extracellular, but distinct
RT intracellular, domain sequences."
RL Cytokine 2:231-237(1990).
[5]
RP SEQUENCE OF 116-461 FROM N.A. AND PARTIAL SEQUENCE.
RX MEDLINE-90349572; PubMed-216946;
RA Heller R.A., Song K., Onasch W.A., Fischer W.H., Chang D.,
RA Ringold G.M.;
RT "Complementary DNA cloning of a receptor for tumor necrosis factor
RT and demonstration of a shed form of the receptor."
RL Proc. Natl. Acad. Sci. U.S.A. 87:6151-6155(1990).
[6]
RP SEQUENCE OF 27-31.
RX MEDLINE-90110215; PubMed-2153136;
RA Engelmann H., Novick D., Wallach D.;
RT "Two tumor necrosis factor-binding proteins purified from human
RT urine. Evidence for immunological cross-reactivity with cell surface
RT tumor necrosis factor receptors."
RL J. Biol. Chem. 265:1531-1536(1990).
[7]
RP SEQUENCE OF 23-40; 65-69; 136-141; 300-306 AND 346-362.
RX MEDLINE-91056048; PubMed-2173696;
RA Loetscher H., Schlaeger E.J., Lahm H.W., Pan Y.-C.E., Lesslauer W.,
RA Brockhaus M.;
RT "Purification and partial amino acid sequence analysis of two
RT distinct tumor necrosis factor receptors from H160 cells."
RL J. Biol. Chem. 265:20131-20138(1990).
[8]
RP CHARACTERIZATION.
RX MEDLINE-93016040; PubMed-1328224;
RA Pennica D., Lam V.T., Mize N.K., Weber R.F., Lewis M., Fendly B.M.,
RA Lipari M.T., Goeddel D.V.;
RT "Biochemical properties of the 75-kDa tumor necrosis factor receptor.
RT Characterization of ligand binding, internalization, and receptor
RT phosphorylation."
RL J. Biol. Chem. 267:21172-21178(1992).
[9]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 419-428 IN COMPLEX WITH
RP TNF2.
RX MEDLINE-99221490; PubMed-10206649;
RA Park Y.C., Burklit V., Villa A.R., Tong L., Wu H.;
RT "Structural basis for self-association and receptor recognition of
RT human TNF2."
RL Nature 398:533-538(1999).
[10]
RP FUNCTION: Receptor with high affinity for TNFSF2/TNF-alpha and
RP approximately 5-fold lower affinity for homotrimeric
RP TNFSF1/lymphotoxin-alpha.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein and secreted.
CC -1- PTM: Phosphorylated; mainly on serine residues and with a very low
CC level on threonine residues.

```


CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch)

DR EMBL: L04270; AAA6757.1; -
 DR EMBL: BC026262; AAA26262.1; -
 DR HSSP: P25942; 1CDF.
 DR Genew: HGNC:6718; LTBR.

DR InterPro: IPR001368; TNFR_c6.
 DR Pfam: PF00020; TNFR_c6; 4.
 DR ProDom: PD000771; TNFR_c6; 1.
 DR SMART: SM00208; TNFR; 4.

DR PROSITE: PS00652; TNFR_NGFR_1; 2.
 DR PROSITE: PS00652; TNFR_NGFR_2; 3.
 DR Receptor: Apoptosis; Transmembrane; Glycoprotein; Repeat; Signal.

FT CHAIN 1 30
 FT SIGNAL 31 435
 FT POTENTIAL.

FT TUMOR NECROSIS FACTOR RECEPTOR
 FT SUPERFAMILY MEMBER 3.
 FT EXTRACELLULAR (POTENTIAL).

FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).

FT REPEAT 42 81
 FT REPEAT 82 124
 FT REPEAT 125 168
 FT REPEAT 169 211
 FT REPEAT 211 254

FT DISULFID 43 58
 FT DISULFID 59 72
 FT DISULFID 62 80
 FT DISULFID 83 98
 FT DISULFID 101 116
 FT DISULFID 104 124
 FT DISULFID 126 132
 FT DISULFID 139 148
 FT DISULFID 142 167
 FT DISULFID 170 185
 FT CARBOHYD 40 40
 FT CARBOHYD 177 177

FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).

SO SEQUENCE 435 AA; 46709 MW; 6246266022F656F CRC64;

Query Match 98.8%; Score 2220; DB 1; Length 435;
 Best Local Similarity 91.7%; Pred. No. 7.1e-120;
 Matches 399; Conservative 0; Mismatches 0; Indels 36; Gaps 1;

QY 1 MLPMATISAPGLAMGLVLTGLFGLAASQPAVPPYASENQTGRDQKEYYEPOHRIICS 60
 DB 1 MLPMATISAPGLAMGLVLTGLFGLAASQPAVPPYASENQTGRDQKEYYEPOHRIICS 60
 QY RCPGGTYVSAKSGRIDTYCATCAENSYNHNWYLTICQLCRPCDPVYGLLETAICTSKR 120
 DB RCPGGTYVSAKSGRIDTYCATCAENSYNHNWYLTICQLCRPCDPVYGLLETAICTSKR 120
 QY 121 KTCRCQOPGFCAMALECTHCELLSDCPGTAELKDEYKGNNNHCVPKKAHPONTSS 180
 DB 121 KTCRCQOPGFCAMALECTHCELLSDCPGTAELKDEYKGNNNHCVPKKAHPONTSS 180
 QY 181 PSARQCPHRCENQGLVEAPGTAOSTTCKNLEPLPEMS-----222
 DB 181 PSARQCPHRCENQGLVEAPGTAOSTTCKNLEPLPEMS-----222
 QY 223 -----GSLKRRPQGGGPNPVAGSMPPRAHPYFPDLPVPLPISGD 264
 DB 223 -----GSLKRRPQGGGPNPVAGSMPPRAHPYFPDLPVPLPISGD 264
 QY 265 VSPVSTGLPAAPVLAGVPOQOSPLDLTREPQLEPGESQVAGHTNGIHVTGSGMTTGN 324
 DB 265 VSPVSTGLPAAPVLAGVPOQOSPLDLTREPQLEPGESQVAGHTNGIHVTGSGMTTGN 324
 QY 301 VSPVSTGLPAAPVLAGVPOQOSPLDLTREPQLEPGESQVAGHTNGIHVTGSGMTTGN 360
 DB 301 VSPVSTGLPAAPVLAGVPOQOSPLDLTREPQLEPGESQVAGHTNGIHVTGSGMTTGN 360

QY 325 IYINPVLGPPGPDLPATPEPPYPIPEEGDPGPGISTPOHDEGKAMHIAETRHCA 384
 DB 361 IYINPVLGPPGPDLPATPEPPYPIPEEGDPGPGISTPOHDEGKAMHIAETRHCA 420
 QY 385 TTSNCGPRNOFTTHD 399
 DB 421 TTSNCGPRNOFTTHD 435

RESULT 2

TNR3_MOUSE
 ID TNR3_MOUSE STANDARD: PRI; 415 AA.
 AC P50284;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tumor necrosis factor receptor superfamily member 3 precursor
 DE (Lymphotoxin-beta receptor).
 GN LTBR OR TNFRSF3 OR TNFR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=CVB; TISSUE=Lung;
 RX MEDLINE=96072804; PubMed=7594541;
 RA Force W.R., Walter B.N., Hession C., Tizard R., Kozak C.A.,
 RA Browning J.L., Ware C.F.,
 RT "Mouse lymphotoxin-beta receptor. Molecular genetics, ligand binding,
 RT and expression.";
 RT J. Immunol. 155:5280-5288(1995).
 RL [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96163885; PubMed=8586432;
 RA Nakamura T., Tashiro K., Nazarea M., Nakano T., Sasayama S.,
 RA Honjo T.;
 RT "The murine lymphotoxin-beta receptor cDNA: isolation by the signal
 RT sequence trap and chromosomal mapping.";
 RL Genomics 30:312-319(1995).
 CC CC
 CC -1- FUNCTION: Receptor for the heterotrimeric lymphotoxin containing
 CC LTA and LTb, and for TNFS14/LIGHT. Promotes apoptosis via TRAF3
 CC and TRAF5. May play a role in the development of lymphoid organs
 CC (By similarity).
 CC -1- SUBUNIT: Self-associates (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.
 CC CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch)

DR EMBL: U29173; AAA6964.1; -
 DR EMBL: U38423; AAB00846.1; -
 DR EMBL: U30798; AAB1334.1; -
 DR HSSP: O14763; 1DOG.
 DR MGD: MGI:104875; Ltbr.
 DR InterPro: IPR001368; TNFR_c6.
 DR Pfam: PF00020; TNFR_c6; 3.
 DR ProDom: PD000771; TNFR_c6; 1.
 DR SMART: SM00208; TNFR; 3.
 DR PROSITE: PS00652; TNFR_NGFR_1; 2.
 DR PROSITE: PS00652; TNFR_NGFR_2; 3.
 DR Receptor: Apoptosis; Transmembrane; Glycoprotein; Repeat; Signal.
 FT CHAIN 1 30
 FT SIGNAL 31 415
 FT POTENTIAL.
 FT TUMOR NECROSIS FACTOR RECEPTOR
 FT SUPERFAMILY MEMBER 3.
 FT EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 224 244

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2003, 10:39:30 ; Search time 27.6158 Seconds

(Without alignments)
\$99.261 Million cell updates/sec

Title: US-09-917-372-1

Perfect score: 2248

Sequence: 1 MLIPWATSAPGLAWGPIVLG.....EHGCGTSPNRCPRNQFTIHD 399

Scoring table: BIOSUM62

Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 405 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2220	98.8	435	1	TNR3_HUMAN
2	1404	62.5	415	1	TNR3_MOUSE
3	349.5	15.5	461	1	TR1B_HUMAN
4	327.5	14.6	474	1	TR1B_MOUSE
5	319	14.2	300	1	TR6B_HUMAN
6	314.5	14.0	277	1	TNR5_HUMAN
7	312	13.9	625	1	TR1_MOUSE
8	298	13.3	616	1	TR1_HUMAN
9	253	11.3	289	1	TNR5_MOUSE
10	251.5	11.2	269	1	TNR5_BOVIN
11	244.5	10.9	461	1	TR1A_PIG
12	244	10.9	401	1	T11B_MOUSE
13	243	10.8	655	1	TR1_MOUSE
14	240	10.7	283	1	TR14_HUMAN
15	239.5	10.7	401	1	T11B_HUMAN
16	239	10.6	595	1	TNR8_HUMAN
17	239	10.6	655	1	TR21_HUMAN
18	235.5	10.5	401	1	T11B_RAT
19	232	10.3	271	1	TNR4_RAT
20	231	10.3	471	1	TR1A_BOVIN
21	228	10.1	454	1	TR1A_MOUSE
22	226	10.1	272	1	TNR4_MOUSE
23	225	10.0	349	1	CRMB_CAMPS
24	223	9.9	498	1	TNR8_MOUSE
25	222.5	9.9	455	1	TR1A_HUMAN
26	212	9.4	323	1	TNR6_BOVIN
27	210	9.3	349	1	CRMB_VARV
28	209.5	9.3	277	1	TNR4_HUMAN
29	207	9.2	351	1	CRMB_COMPA
30	204	9.1	461	1	TR1A_RAT
31	203.5	9.1	255	1	TNR9_HUMAN
32	196.5	8.7	425	1	TR16_RAT
33	194	8.6	332	1	TNR6_PIG

34	191.5	8.5	493	1	TNR8_RAT	P97525 rattus norv
35	189	8.4	325	1	VT2_SEVKA	P25943 shope fibro
36	185.5	8.3	256	1	TNR9_MOUSE	P20334 mus musculu
37	185	8.2	440	1	T10B_HUMAN	O14763 homo sapien
38	183.5	8.2	326	1	VT2_MITXV	P23825 myxoma viru
39	183.5	8.2	335	1	TNR6_HUMAN	P23445 homo sapien
40	183.5	8.2	416	1	TR16_CHICK	P18519 gallus galli
41	175.5	7.8	417	1	TR16_MOUSE	O94041 mus musculu
42	175.5	7.8	427	1	TR12_HUMAN	P08138 homo sapien
43	164	7.3	327	1	TR12_MOUSE	O93038 h tumor nec
44	162.5	7.2	327	1	TNR6_MOUSE	P23446 mus musculu
45	149.5	6.7	180	1	TR22_MOUSE	O9er62 mus musculu

ALIGNMENTS

RESULT 1	ID	TNR3_HUMAN	STANDARD:	PRT:	435 AA.
AC	P36941:	01-JUN-1994 (Rel. 29, Created)			
DT		01-JUN-1994 (Rel. 29, Last sequence update)			
DE		15-JUN-2002 (Rel. 41, Last annotation update)			
DE		Tumor necrosis factor receptor superfamily member 3 precursor			
DE		(Lymphotoxin-beta receptor) (Tumor necrosis factor receptor. 2 related protein) (Tumor necrosis factor C receptor).			
GN		LTBR OR TNFRSF3 OR TNFR.			
OS		Homo sapiens (Human).			
OC		Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC		Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX		NCBI_TaxID=9606;			
RN		[1]			
RP		SEQUENCE FROM N.A.			
RC		TISSUE=Liver;			
RX		MEDLINE=93252381; PubMed=8486360;			
RA		Baens M., Chaffanet M., Cassiman J.J., den Berghe H., Marynen P.;			
RT		"Construction and evaluation of a hncDNA library of human 12p			
RL		transcribed sequences derived from a somatic cell hybrid."			
RL		Genomics 16:214-218(1993).			
RN		[2]			
RP		SEQUENCE FROM N.A.			
RC		TISSUE=Lung;			
RA		Straussberg R.;			
RL		Submitted (Apr-2002) to the EMBL/Genbank/DBJ databases.			
RN		[3]			
RP		FUNCTION.			
RX		MEDLINE=94225209; PubMed=8171323;			
RA		Crowe P.D., van Arsdale T.L., Walter B.N., Ware C.F., Hession C.,			
RT		Ehrenfels B., Browning J.L., Din W.S., Goodwin R.G., Smith C.A.;			
RL		"A lymphotoxin-beta-specific receptor."			
RL		Science 264:707-710(1994).			
RN		[4]			
RP		CHARACTERIZATION.			
RX		MEDLINE=99223511; PubMed=10207006;			
RA		Wu M.-Y., Wang P.-Y., Han S.-H., Hsieh S.-L.;			
RT		"The cytoplasmic domain of the lymphotoxin-beta receptor mediates cell			
RL		death in Hela cells."			
RN		J. Biol. Chem. 274:11868-11873(1999).			
RN		[5]			
RP		FUNCTION.			
RX		MEDLINE=20261554; PubMed=10799510;			
RA		Rooney I.A., Butovich K.D., Glas A.A., Borboroglu S., Benedict C.A.,			
RT		Whitbeck J.C., Cohen G.H., Eisenberg R.D., Ware C.F.;			
RL		"The lymphotoxin-beta receptor is necessary and sufficient for			
RL		light-mediated apoptosis of tumor cells."			
RL		J. Biol. Chem. 275:14307-14315(2000).			
CC		-1- FUNCTION: Receptor for the heterotrimeric lymphotoxin containing			
CC		LTA and LTB, and for TNFSF14/LIGHT. Promotes apoptosis via TRAF3			
CC		and TRAF5. May play a role in the development of lymphoid organs.			
CC		-1- SUBUNIT: Self-associates.			
CC		-1- SUBCELLULAR LOCATION: Type I membrane protein.			
CC		-1- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.			


```

* 93976 98771: contig of 4796 bp in length
* 98772 98871: gap of unknown length
* 98872 101999: contig of 3128 bp in length
* 102000 102009: gap of unknown length
* 102100 105352: contig of 3253 bp in length
* 105353 105452: gap of unknown length
* 105453 109873: contig of 4421 bp in length
* 109874 109973: gap of unknown length
* 109974 112881: contig of 2908 bp in length
* 112882 117279: gap of unknown length
* 117280 117380: gap of unknown length
* 121635 121734: gap of unknown length
* 121735 125920: contig of 4186 bp in length
* 125921 126020: gap of unknown length
* 126021 129909: contig of 3889 bp in length
* 129910 130009: gap of unknown length
* 130010 135723: contig of 5714 bp in length
* 135724 135823: gap of unknown length

```

Query Match
 Best Local Similarity 67.9% Pred. No. 1.8e-43;
 Matches 558; Conservative 0; Mismatches 225; Indels 39; Gaps 12;

```

QY 1162 CCGGGGGGCTGACTGACTGCAACATCTACATCTACATGAGACCACTGAGGG 1221
Db 154396 CTGCAAGGCTGTGACTGCTACCGGCAATCTACATCTACATGAGGCGG 154455
QY 1222 GACCAACGGGCTCTGAGACCTCCGAGTACCCCGGACCTCCATCCCGAAG 1281
Db 154456 GAAACAGGGGGCTGTGAGACCTCCGAGTACCCCGGACCTCCCGAAG 154515
QY 1282 AGGGGAGCCCTGCGCCCTCCGAGCTCTACACACCCAGAGAGATGGAAGGCTTGGC 1341
Db 154516 AGGGAGCCCTGCGCCCTCTGTGAGCTGTACACCTTACAGAGAGATGGAAGCTTGGC 154575
QY 1342 ACCTGGGGAGACAGAG-CAGTGTGTGCTCCACACCTCTTAACAGGGGCCCAAGAACCA 1400
Db 154576 ATCTGGGCTGAGACAGAGACCTGAGGTGCGCAAGACCTCTGAC--GGGGCCAAAGGACCTA 154633
QY 1401 TTTATACCCCAWGACTGAGTGTGAGAAAGGCAAGAAAGGGGGGCAAGAGGCAC 1460
Db 154634 TTAGCCACCCATGCTGATGATCTGTGGAAAGGCAAGAGAGAGAGGCTCA--GCAC 154691
QY 1461 CTCTCCCTTGAAGGCTGCTGCGCCACAGTGGATTTCAAGGGGCTGAGTGAAGGCGCGG 1520
Db 154692 CTCTCTCC-TGAGGCTGCTTACCCACAGAGCTCCAGAG--TTTGGCTGGGCTCTGA 154746
QY 1521 GAACGAGAGCCCTAAGGATTAAGGCTCAGACACCTCTGAGAGCAGGTGGGCACTGGCTG 1580
Db 154747 GATACAGAGCCCTAAGGAGACTTGGCTCAGACACCTCT---TCAGAGAGACACTAGCGG 154802
QY 1581 GGTAGGCTGCTTACAGAGACTTCCCTACTGCTGAGCAAACTGAGGCTCCCGGCA 1640
Db 154803 GGCACGCTGCGCCGCG--CAGGCGCTGAGACACCGCTGAGAGGCTTGTGCA 154861
QY 1641 GACCCACACCCCTTGGGCTGCTCAGCCTCAGGCAAGGAGGAGGACATGATTAACAAC 1700
Db 154862 GACCCACACCTTCCCAAG-----GCTCAGACAGAGCAAGGTTGTGACCCAGC 154913
QY 1701 TGCTGCCCATCAGACAGCGCGCAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1760
Db 154914 ATGCGCTCTGAGAGCCCTAAACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 154965
QY 1761 CTTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1820
Db 154966 CTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 155024
QY 1821 GACCCCTTTGGGCTTCAACACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1880
Db 155025 GGCCCTTGAAGGCTTGTAC-----TGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 155080

```

```

QY 1881 ATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1940
Db 155081 GCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 155137
QY 1941 AGGGGCTTGGTGTAGATTTCTAGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1982
Db 155138 GGGGCTTCAAGTATTAACCAATGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 155179

```

```

RESULT 15
HUMTUMNEC
LOCUS
DEFINITION
Homo sapiens (clone MCD18) tumor necrosis factor receptor related
protein mRNA, complete exon and repeat region.
ACCESSION
L04489
VERSION
L04489.1 GI:340022
KEYWORDS
tumor necrosis factor receptor related protein.
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
Baens,M., Chaffanet,M., Cassiman,J.J., van den Bergh,H. and
Mayren,P.
Construction and evaluation of a hncDNA library of human 12p
transcribed sequences derived from a somatic cell hybrid
genomics 16 (1), 214-218 (1993)
93252381
8486360
FEATURES
source
1..1605
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="12p13"
/cell_line="M28 somatic cell hybrid"
/tissue_id="hncdna tc651"
1..1605
/gene="TNFR"
repeat_region
1..105
/partial
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed
250..444
/gene="TNFR"
/note="L04270 sequence homologue"
/evidence="experimental"
BASE COUNT
472 a 394 c 323 g 416 t
ORIGIN

```

Query Match
 Best Local Similarity 10.0%; Score 198; DB 9; Length 1605;
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 689 AGATGAGTGGGAGAGGTAACACACTGGCTCCCGCAAGAGAGAGAGAGAGAGAGAGAGAG 748
Db 248 AGATGAGTGGGAGAGGTAACACACTGGCTCCCGCAAGAGAGAGAGAGAGAGAGAGAGAG 307
QY 749 TACCTCTCCCGAGCGCCGCTGCGACCCACACAGAGTGTGAGAACCAAGTCTGGT 808
Db 308 TACCTCTCTCCCGAGCGCCGCTGCGACCCACACAGAGTGTGAGAACCAAGTCTGGT 367
QY 809 GAGGACAGCTCCAGGACATGCGCCAGTCCGACACAACTGCAAAAATCCATTAGAGCACT 868
Db 368 GAGGACAGCTCCAGGACATGCGCCAGTCCGACACAACTGCAAAAATCCATTAGAGCACT 427
QY 869 GCGCCCAAGATGTGAGG 886
Db 428 GCGCCCAAGATGTGAGG 445

```

Search completed: April 15, 2003, 15:12:59
 Job time : 4171.99 secs

Lozada, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Mageshwar, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
 Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metker, M.,
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogun, M., Okunolu, G.,
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
 Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G.,
 Scherer, S., Scott, G., Shen, H., Shoostari, N., Slason, I.,
 Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Sytek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Tellford, B., Thomas, R., Thomas, S.,
 Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstein, G., and Gibbs, R.
 Direct Submission
 Unpublished
 2 (bases 1 to 187998)
 Worley, K.C.
 Direct Submission
 Submitted (02-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 187998)
 Worley, K.C.
 Direct Submission
 Submitted (23-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GWSA
 Center clone name: CH230-100N7
 ----- Summary Statistics
 Sequencing vector: Plasmid
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 133860 bases at least Q40
 Consensus quality: 142697 bases at least Q30
 Consensus quality: 148841 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 64 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1 1031: contig of 1031 bp in length
 * 1032 1131: gap of unknown length
 * 1132 2334: contig of 1203 bp in length
 * 2335 2434: gap of unknown length
 * 2435 3488: contig of 1054 bp in length
 * 3489 3588: gap of unknown length
 * 3589 4860: contig of 1272 bp in length
 * 4861 4960: gap of unknown length
 * 4961 6021: contig of 1061 bp in length
 * 6022 6121: gap of unknown length
 * 6122 7174: contig of 1053 bp in length
 * 7175 7274: gap of unknown length
 * 7275 8532: contig of 1258 bp in length
 * 8533 8632: gap of unknown length
 * 8633 9711: contig of 1079 bp in length
 *
 * 9712 9811: gap of unknown length
 * 9812 11308: contig of 1497 bp in length
 * 11309 12604: gap of unknown length
 * 11409 12604: contig of 1196 bp in length
 * 12605 12704: gap of unknown length
 * 12705 14114: contig of 1410 bp in length
 * 14115 14214: gap of unknown length
 * 14215 16707: contig of 2492 bp in length
 * 16707 16807: gap of unknown length
 * 16807 17876: contig of 1069 bp in length
 * 17876 17975: gap of unknown length
 * 17975 19654: contig of 1679 bp in length
 * 19655 19754: gap of unknown length
 * 19755 21684: contig of 1929 bp in length
 * 21684 21783: gap of unknown length
 * 21783 22871: contig of 1088 bp in length
 * 22871 22972: gap of unknown length
 * 22972 24667: contig of 1695 bp in length
 * 24667 26943: gap of unknown length
 * 26943 27043: contig of 2177 bp in length
 * 27043 29057: gap of unknown length
 * 29057 29157: contig of 2014 bp in length
 * 29157 30973: gap of unknown length
 * 30973 31073: contig of 1816 bp in length
 * 31073 32838: gap of unknown length
 * 32838 32938: contig of 1765 bp in length
 * 32938 34616: gap of unknown length
 * 34616 34716: contig of 1678 bp in length
 * 34716 37198: gap of unknown length
 * 37198 37298: contig of 2482 bp in length
 * 37298 39477: gap of unknown length
 * 39477 39577: contig of 2179 bp in length
 * 39577 41815: gap of unknown length
 * 41815 41915: contig of 2238 bp in length
 * 41915 44488: gap of unknown length
 * 44488 44588: contig of 2573 bp in length
 * 44588 46062: gap of unknown length
 * 46062 46162: contig of 1474 bp in length
 * 46162 48328: gap of unknown length
 * 48328 48428: contig of 2166 bp in length
 * 48428 51393: gap of unknown length
 * 51393 51493: contig of 2367 bp in length
 * 51493 54159: gap of unknown length
 * 54159 54259: contig of 2664 bp in length
 * 54259 56383: gap of unknown length
 * 56383 56483: contig of 2124 bp in length
 * 56483 58779: gap of unknown length
 * 58779 60911: contig of 2296 bp in length
 * 60911 61011: gap of unknown length
 * 61011 63907: contig of 2032 bp in length
 * 63907 64007: gap of unknown length
 * 64007 66202: contig of 2896 bp in length
 * 66202 66302: gap of unknown length
 * 66302 69738: contig of 2195 bp in length
 * 69738 69838: gap of unknown length
 * 69838 72410: contig of 3436 bp in length
 * 72410 72510: gap of unknown length
 * 72510 76219: gap of unknown length
 * 76219 76319: contig of 3709 bp in length
 * 76319 79029: gap of unknown length
 * 79029 79129: contig of 2710 bp in length
 * 79129 83311: gap of unknown length
 * 83311 83411: contig of 4182 bp in length
 * 83411 86394: gap of unknown length
 * 86394 86494: contig of 2983 bp in length
 * 86494 88582: gap of unknown length
 * 88582 90972: contig of 2088 bp in length
 * 90972 91072: gap of unknown length
 * 91072 93875: contig of 2290 bp in length
 * 93875 93975: gap of unknown length
 * 93975 93876: gap of unknown length

```

* 53463 55170: contig of 1708 bp in length
* 55171 55270: gap of unknown length
* 55271 55697: contig of 1427 bp in length
* 56698 56797: gap of unknown length
* 56798 58037: contig of 1240 bp in length
* 58038 58137: gap of unknown length
* 60362 60461: contig of 2224 bp in length
* 60462 62335: gap of unknown length
* 62336 62435: contig of 1874 bp in length
* 62436 64906: gap of unknown length
* 64907 65006: contig of 2471 bp in length
* 65007 67476: gap of unknown length
* 67476 67575: contig of 2469 bp in length
* 67576 69653: gap of unknown length
* 69654 69753: contig of 2078 bp in length
* 69754 73830: gap of unknown length
* 73831 73930: contig of 4077 bp in length
* 73931 77047: gap of unknown length
* 77048 77147: contig of 3117 bp in length
* 77148 80438: gap of unknown length
* 80439 80538: contig of 3291 bp in length
* 80539 83532: gap of unknown length
* 83533 83632: contig of 2894 bp in length
* 83633 85754: gap of unknown length
* 85755 85854: contig of 2122 bp in length
* 85855 89203: gap of unknown length
* 89204 89303: contig of 3349 bp in length
* 89304 92865: gap of unknown length
* 92866 92965: contig of 3562 bp in length
* 92966 95763: gap of unknown length
* 95764 95863: contig of 2798 bp in length
* 95864 99698: gap of unknown length
* 99699 99798: contig of 3835 bp in length
* 99799 104352: gap of unknown length
* 104353 104452: contig of 4554 bp in length
* 104453 108165: gap of unknown length
* 108166 108265: contig of 3713 bp in length
* 108266 111945: gap of unknown length
* 111946 112045: contig of 3680 bp in length
* 112046 116301: gap of unknown length
* 116302 116401: contig of 4256 bp in length
* 116402 120825: gap of unknown length
* 120826 120925: contig of 4424 bp in length

```

Query Match 14.28; Score 281.8; DB 2; Length 178228;
 Best Local Similarity 68.28; Pred. No. 4.6e-47;
 Matches 579; Conservative 0; Mismatches 232; Indels 38; Gaps 12;

```

QY 1135 CCCACGGTACCAATGGCATTCATGTCAACGGGGCTGTATGACTATCATCTGGCAACATCT 1194
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 152113 CCACAGGTGGATGGATTGACATTCAGTCGACGAGGCTCTGTGACTGTACCGCAATATCT 152054

QY 1195 ACATCTCAATATGACACAGTACTGTGGGGGACACCGGGTCTGTGAGACCTCCACAGTACC 1254
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 152053 ACATATCAATATGACACAGTACTGTGGGGGACACCGGGTCTGTGAGACCTCCACAGTACC 151994

QY 1255 CCGAGACTCCATACCCCATTCGCCGAAGAGGGGAGACCTGGCCCTCCCGGCTCTTACAC 1314
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151993 CTGAGCTTCATACCCCATTCGCCGAAGAGGGGAGACCTGGCCCTCCCGGCTCTTACAC 151934

QY 1315 CCCACGAGAAATGGCAAGCTTGGCACTACCGAGAGACAGAG-CACGTGTGGTCCACA 1373
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151933 CCAACGAGAGAGATGGCAAGCTTGGCACTACCGAGAGACAGACATAGAGGTGCCAAG 151874

QY 1374 CCTCTTACAGGGGGCCCAAGAACCAATTATATGACATGACATGAGTGTGAGAAAA 1433
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151873 ACCTCTGAC--GGGGCCCAAGGACATATAGACACCAATGCTGAGCATCTGGAAAA 151816

QY 1434 GGCAGAAAGAGGGGGGACAAAGGACCTTCTCCCTTGGAGCTGCCCTCCACAGTGGGA 1493
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151815 GCAAGAGAGAGAGGGGTCCA--GCACCTCTTCC--TGAGGCTGCTCTACCAACAGG-- 151761

```

```

QY 1494 TTCACAGGGGCTGTAGTAGAGGCCCGGAGACAGAGCCCTTAAGGATTTAAGGCTACAGACA 1553
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151760 -ACTACAGAGATTGGTGGCTGGGCTGTGAGATACAGAGCCCTTAAGGACTTGGCTCAGACA 151702

QY 1554 CCTCTGAGAGACAGTGGGACACTGTGGGTAGAGGTGTGCCCTCCACAGACTCTCCCTACTG 1613
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151701 CCTCT---TGAGAGAGACATACAGCGGGGACCGCTGCCCGCC--CAGGCCCTGTACACCG 151647

QY 1614 CCGTGAACCTGTAGGCGCTCCGGAGAGACCCACCCACCCCTTGGGCTGTACAGCTCA 1673
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151646 CCTGAGAGAGCTTGTAGAGCTTGTTCAGACACCCACACTCCCAAG-----GCTCA 151595

QY 1674 GGCACGAGACAGGACATATGATACCACTGTGCCCTACAGACAGCGGACCGAGACAC 1733
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151594 GACACAGGACAGAGCTGTGTGACCCACCATGCGCTCTGTAGACCCCTTAACATAGGACAGAG 151535

QY 1734 GGCACGAGAGAGAGCGCCGACACAGGTACCTGACATGACAGACGTCAGGGCCCTTAAGGAT 1793
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151534 GGAAGCTGG-----CCAGGGCTCACTTCAAGGACATC--CGGGGACCTTGCAAGGCT 151484

QY 1794 TCGTGTGCTCATCCCAACCTTCAGAGACCTTTGGGGTTCCACACTTCACGTGACATG 1853
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151483 TTGTGTGCTCATCCCAACCTTCAGAGACCTTTGGGGTTCCACACTTCACGTGACATG 151428

QY 1854 AGGTAGACCTTCATGATGAATGAATTAAGGAGAGAGCGCTCCCTCCCTCCCTAGA 1913
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151427 AGATAGACCTTCATGATGAATGAATGAATGAAGAGAGAGCGCTCCCTCCCTCCCTAG 151371

QY 1914 GGAAGAGAGAGAGAGATTAACACTAGAGGGGTTGGGTAGATTTCTTCTAGATAGGAGAA 1973
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151370 GGGAGAGAGAGAGAGATTAACACTAGAGGGGTTGGGTAGATTTCTTCTAGATAGGAGAA 151311

QY 1974 GAGTTTGG 1982
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151310 CAGTTTGG 151302

```

RESULT 14
 AC125909
 LOCUS
 DEFINITION
 Rattus norvegicus clone CH230-100N7, *** SEQUENCING IN PROGRESS
 *** 64 unordered pieces.
 ACCESSION
 AC125909
 VERSION
 AC125909.1 GI:21671507
 KEYWORDS
 HTG: HTGS, PHASE1.
 SOURCE
 Rattus norvegicus
 ORGANISM
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE
 AUTHORS
 1 (bases 1 to 187998)
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
 Alibrooks,S.B., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,
 Barbara,J., Benton,J., Bimaye,K., Blankenburg,K., Bonini,D.,
 Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
 Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dahorne,S.R., David,R.,
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
 Douhaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escoto,M.,
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
 Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
 Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
 Homsl,F., Howard,S., Huber,J., Huix,S., Hume,J., Jackson,L.E.,
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
 Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
 Li,J., Li,Z., Lichteage,O., Lieu,C., Liu,C., Liu,J., Liu,W., Lounseged,H.,

Denaturation:	94 degrees C for 15 seconds
Annealing:	62 degrees C for 23 seconds

Polymerization:	72 degrees C for 30 seconds
PCR cycles:	30

Thermal Cycler: Perkin Elmer 9600

Template:	25 ng
Primer:	each 1 μ M

dNTPs: each 200 μ M
Taq Polymerase: 0.05 units/ μ l

Total Vol: 10 ul

MgCl₂: 2.5 ml

KCl:	50	ml
Tris-HCl:	20	ml

3.8
:H \bar{p}

Location/Qualifiers

```

1: 400
/organism="Homo sapiens"

```

```

/obj/alel-
/axon:9606-
/map="12"

```

89. .220

```
complement(202, .220)
```

[illegible]

15.0%; Score 296.8; DB 11; Length 400;
97.0%; Prod 8.0; FA 0.0; *W* 0.0; *W*₁ 0.0; *W*₂ 0.0; *W*₃ 0.0; *W*₄ 0.0; *W*₅ 0.0; *W*₆ 0.0; *W*₇ 0.0; *W*₈ 0.0; *W*₉ 0.0; *W*₁₀ 0.0; *W*₁₁ 0.0; *W*₁₂ 0.0; *W*₁₃ 0.0; *W*₁₄ 0.0; *W*₁₅ 0.0; *W*₁₆ 0.0; *W*₁₇ 0.0; *W*₁₈ 0.0; *W*₁₉ 0.0; *W*₂₀ 0.0; *W*₂₁ 0.0; *W*₂₂ 0.0; *W*₂₃ 0.0; *W*₂₄ 0.0; *W*₂₅ 0.0; *W*₂₆ 0.0; *W*₂₇ 0.0; *W*₂₈ 0.0; *W*₂₉ 0.0; *W*₃₀ 0.0; *W*₃₁ 0.0; *W*₃₂ 0.0; *W*₃₃ 0.0; *W*₃₄ 0.0; *W*₃₅ 0.0; *W*₃₆ 0.0; *W*₃₇ 0.0; *W*₃₈ 0.0; *W*₃₉ 0.0; *W*₄₀ 0.0; *W*₄₁ 0.0; *W*₄₂ 0.0; *W*₄₃ 0.0; *W*₄₄ 0.0; *W*₄₅ 0.0; *W*₄₆ 0.0; *W*₄₇ 0.0; *W*₄₈ 0.0; *W*₄₉ 0.0; *W*₅₀ 0.0; *W*₅₁ 0.0; *W*₅₂ 0.0; *W*₅₃ 0.0; *W*₅₄ 0.0; *W*₅₅ 0.0; *W*₅₆ 0.0; *W*₅₇ 0.0; *W*₅₈ 0.0; *W*₅₉ 0.0; *W*₆₀ 0.0; *W*₆₁ 0.0; *W*₆₂ 0.0; *W*₆₃ 0.0; *W*₆₄ 0.0; *W*₆₅ 0.0; *W*₆₆ 0.0; *W*₆₇ 0.0; *W*₆₈ 0.0; *W*₆₉ 0.0; *W*₇₀ 0.0; *W*₇₁ 0.0; *W*₇₂ 0.0; *W*₇₃ 0.0; *W*₇₄ 0.0; *W*₇₅ 0.0; *W*₇₆ 0.0; *W*₇₇ 0.0; *W*₇₈ 0.0; *W*₇₉ 0.0; *W*₈₀ 0.0; *W*₈₁ 0.0; *W*₈₂ 0.0; *W*₈₃ 0.0; *W*₈₄ 0.0; *W*₈₅ 0.0; *W*₈₆ 0.0; *W*₈₇ 0.0; *W*₈₈ 0.0; *W*₈₉ 0.0; *W*₉₀ 0.0; *W*₉₁ 0.0; *W*₉₂ 0.0; *W*₉₃ 0.0; *W*₉₄ 0.0; *W*₉₅ 0.0; *W*₉₆ 0.0; *W*₉₇ 0.0; *W*₉₈ 0.0; *W*₉₉ 0.0; *W*₁₀₀ 0.0; *W*₁₀₁ 0.0; *W*₁₀₂ 0.0; *W*₁₀₃ 0.0; *W*₁₀₄ 0.0; *W*₁₀₅ 0.0; *W*₁₀₆ 0.0; *W*₁₀₇ 0.0; *W*₁₀₈ 0.0; *W*₁₀₉ 0.0; *W*₁₁₀ 0.0; *W*₁₁₁ 0.0; *W*₁₁₂ 0.0; *W*₁₁₃ 0.0; *W*₁₁₄ 0.0; *W*₁₁₅ 0.0; *W*₁₁₆ 0.0; *W*₁₁₇ 0.0; *W*₁₁₈ 0.0; *W*₁₁₉ 0.0; *W*₁₂₀ 0.0; *W*₁₂₁ 0.0; *W*₁₂₂ 0.0; *W*₁₂₃ 0.0; *W*₁₂₄ 0.0; *W*₁₂₅ 0.0; *W*₁₂₆ 0.0; *W*₁₂₇ 0.0; *W*₁₂₈ 0.0; *W*₁₂₉ 0.0; *W*₁₃₀ 0.0; *W*₁₃₁ 0.0; *W*₁₃₂ 0.0; *W*₁₃₃ 0.0; *W*₁₃₄ 0.0; *W*₁₃₅ 0.0; *W*₁₃₆ 0.0; *W*₁₃₇ 0.0; *W*₁₃₈ 0.0; *W*₁₃₉ 0.0; *W*₁₄₀ 0.0; *W*₁₄₁ 0.0; *W*₁₄₂ 0.0; *W*₁₄₃ 0.0; *W*₁₄₄ 0.0; *W*₁₄₅ 0.0; *W*₁₄₆ 0.0; *W*₁₄₇ 0.0; *W*₁₄₈ 0.0; *W*₁₄₉ 0.0; *W*₁₅₀ 0.0; *W*₁₅₁ 0.0; *W*₁₅₂ 0.0; *W*₁₅₃ 0.0; *W*₁₅₄ 0.0; *W*₁₅₅ 0.0; *W*₁₅₆ 0.0; *W*₁₅₇ 0.0; *W*₁₅₈ 0.0; *W*₁₅₉ 0.0; *W*₁₆₀ 0.0; *W*₁₆₁ 0.0; *W*₁₆₂ 0.0; *W*₁₆₃ 0.0; *W*₁₆₄ 0.0; *W*₁₆₅ 0.0; *W*₁₆₆ 0.0; *W*₁₆₇ 0.0; *W*₁₆₈ 0.0; *W*₁₆₉ 0.0; *W*₁₇₀ 0.0; *W*₁₇₁ 0.0; *W*₁₇₂ 0.0; *W*₁₇₃ 0.0; *W*₁₇₄ 0.0; *W*₁₇₅ 0.0; *W*₁₇₆ 0.0; *W*₁₇₇ 0.0; *W*₁₇₈ 0.0; *W*₁₇₉ 0.0; *W*₁₈₀ 0.0; *W*₁₈₁ 0.0; *W*₁₈₂ 0.0; *W*₁₈₃ 0.0; *W*₁₈₄ 0.0; *W*₁₈₅ 0.0; *W*₁₈₆ 0.0; *W*₁₈₇ 0.0; *W*₁₈₈ 0.0; *W*₁₈₉ 0.0; *W*₁₉₀ 0.0; *W*₁₉₁ 0.0; *W*₁₉₂ 0.0; *W*₁₉₃ 0.0; *W*₁₉₄ 0.0; *W*₁₉₅ 0.0; *W*₁₉₆ 0.0; *W*₁₉₇ 0.0; *W*₁₉₈ 0.0; *W*₁₉₉ 0.0; *W*₂₀₀ 0.0; *W*₂₀₁ 0.0; *W*₂₀₂ 0.0; *W*₂₀₃ 0.0; *W*₂₀₄ 0.0; *W*₂₀₅ 0.0; *W*₂₀₆ 0.0; *W*₂₀₇ 0.0; *W*<

servative	0;	Mismatches	8;	Indels	2;	Gaps	2;
-----------	----	------------	----	--------	----	------	----

TTGCTCAGCCTCAGGCACGACAGGGCACATGATA-CCAACTGCTG-CCCA 1709

TTGCTCACCCCTCAGGCACGGACAGGGGCACATGATACCCACTGTGCCCA 341

CCGACCCGGAGCACCGCACCCGAGGGAGCCGCCACACGGTCACTGCAAG 1769

CCGCACCCGGAGCACGGCACCCGAGGAGCCGCCACACGCTCACCTGCAAG 281

CCCCCTAAAGGATTGTTGGTCTCATCCCCAAGCTTCAGAGACCCCTTG 1829

CCCCCAAGGATTCGTGCTCATCCCCCAAGCTTCAGANGACCTTTC 221

CTTACGCGACCGAGGAGACCCCTGCATGAGATGAATTATAGGAGC 1889

CTATCAGCGGGACAGGAGCGAGGACCCCTCCATGAAGATGAATTTATAGGAGG 161

CCGCGCCTGGTGGAGGACATCAACCAACTAGGGGTG 1949

[illegible][illegible]

178228 bp DNA HTG 19-JUL-1994

unordered pieces.

1 GI:21908679
PHASE1

rvegicus.
rvegicus

; Metazoa; Chordata; Vertebrata; Euteleostomi
Eutheria; Rodentia; Sciurognathi. Muridae. Murinae.

PCDIVLGEFEVAPCTSDRKACRCOPGMSVYLNDNEVHCHEERLVLQPGTEAEVTD
 EIMDYNVCVKRKHFOGTSPPARACRHPHLCRIGLILKRRHGESESPCAPRAD
 GAMLILILSLVLEFETVTLACAMRHPLRCLRIGLILKRRHGESESPCAPRAD
 PHRPDLAEPILPMSGDLSPRSPAPRLPALSLEYVLOOQSPVQARELEAEGBHQVA
 HGANGIHVTSVITGNITITNGPVLGRTGPDPAPEPPYPTPEBAGPSELS
 TPYEDGAMHLAETETLGCODL*
 BASE COUNT 316 a 537 c 455 g 306 t
 ORIGIN
 Query Match 34.0% Score 673.8 DB 10; Length 1614;
 Best Local Similarity 69.5% Pred. No. 4.7e-126;
 Matches 1111; Conservative 0; Mismatches 377; Indels 111; Gaps 10;

48 GGAGCCCTGGAGGCCCGGCGCTGCGCTCCGCGCTGGGGTGCACATCGCCCTGAGTC 107
 25 GGAGCCCTGGAGGCCCGGCGCTCCGCGCGCGCGCGCGGGAAGCTGG----- 75
 108 CCGTCCCAAGGCTCTGGGCTCGGGCAGCGCCGCGCACCGCTGCCAGAGACT-CGGGCTC 166
 76 ----CCAGCCCGCAGGTTTGGGAAGTCGTGGCAACGCTGCTCAGGAGCTCGGGCTTC 131
 167 CTGCTCTCTCCAGGCCCGCA--CGTTGCTGGCGCGCTGGCGAGTGGCGCGCATGCTC 224
 132 CCACCTTCTCTAGACTCCACCGGCTGTGTCAGCGCAGCGAAGCGCCCATGCGC 191
 225 CTGCTTGGGCGACCTCTGCCCCCGGCTGGCTGGGGCTCTGTGGTGGGCTCTTC 284
 192 CTGCCCCGGGCTCTCTCTCTGCGGCTGGCTGGGGCGCACTCCTGTGGAGCTCAGC 251
 285 GGGCTCTTGGCAGCATCGACGCCAGCGGGTGGCTCCATATGCTGGAGAACACAGACC 344
 252 GGGCTTGGGCTCTCTGAGCCAGCGGTGGCGCCCTTATCGCATAAACAGAGACT 311
 345 TGCAGGAGACAGAAAGATATGATGAGCCCGACAGCCGATGCTGCTGCTCCGCTGC 404
 312 TCTGTGGAGACAGACAAATATCTACGAGCCCATGACAGAGCTGTGCTGCTCCGCT 371
 405 CCGCAGGACCTATGTCTCAGCTAAATGTAGCCGATCGGAGACAGATTTGTGCCACA 464
 372 CCCCAGGAGGATTTGTCTTCTGCGTATGACACCGGACAGACAGCGTTTGAAGACT 431
 465 TGTGCGAATTTCTACAGAGGAGGACCTGAACTACCTGACATCGGACGAGTGGCGCG 524
 432 TCCCCCAATATTTCTATATATGACACTGGAACCACTCTCCACCTGCGGAGCTGCGC 491
 525 CCTGTGACCCAGATGAGGCTCGAGAGATTTGCCCTGACAGCAAGCAAGGAGGAGAC 584
 492 CCTGTGACATTTGTGCTGGGCTTTGAGGAGTTGCCCTTGCACCAAGCATGAGGAAAGC 551
 585 CAGTCCGCTGCGAGCGGGAATGTCTGTGCTGCTGGGCGCTCGATGTACACACTG- 643
 552 GAGTCCGCTGTACAGCGGAGGATGTCTGTGATCTGGAATGAGTGTGCACTGT 611
 644 ----CGACTACTTTCTACGCGCGCTGCGACAGGAGGAGGAGGAGGAGGAGGAGG 698
 612 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 671
 699 GGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 758
 672 ATGATATCTAGAGCTAGCTGTCTCCCTTAAGCGGAGCACTTCCAGAACACTTCTCC 731
 732 CTTGAGCGCGCTGTACACCCATACAGATGTGAGATCCAGGCGCTGTGTGAGGAGCT 791
 759 CCGAGCGCGCTGTGAGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 818
 733 CTTGAGCGCGCTGTACACCCATACAGATGTGAGATCCAGGCGCTGTGTGAGGAGCT 791
 819 CAGGAGCTGCGGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 878
 792 CAGGAGCTGCTACTCTGAGATCACTGTGAAAAATCCCGAGAGGAGGAGGAGGAGGAG 851
 879 ATGTAGAGATCGCTCTCAG-----AGGCGTCC 907
 852 CTGCTGGCAGATCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 911

908 GCAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 948
 912 GCGTGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 971
 949 -----CTCGAAGGCCCATTCATCTTCCCTGACTTG 980
 972 GAGGAGAGAAATCTCCCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1031
 981 GTACAGCCACTGCTACCATTTCTGGAGATGTTTCCCGAGTATTCAGAGGCTCCCGGAG 1040
 1032 GCAGAGCCACTCTCTCCATCTGTGAGACTTGTCCCATTCCTCTGCTGAGACCCCGAGAG 1091
 1041 GCCCAGATTTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1100
 1092 GCCCCTCTTGGAGAAATGAGTGTACAAACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1151
 1101 CCGCAGTTGGAACCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1160
 1152 CTGAGAGCTGAACCTGGGAGACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1211
 1161 ACCGCGGCTTATGACTATCTGCAACATCTACATCTACATGAGGAGGAGGAGGAGGAGGAG 1220
 1212 ACCGCGGCTCTGTGACTGTACCGGCAACATCTACATCTACATGAGGAGGAGGAGGAGGAG 1271
 1221 GGACCAACCGGCTCTGAGAGACCTCCAGCTACCCCGAAGCTCCATACCCCATCCGAA 1280
 1272 GGAACAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1331
 1281 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1340
 1332 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1391
 1341 CACCTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1399
 1392 CATCTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1449
 1400 ATTATACCCCATGACTGACTGAGTGTGAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1459
 1450 ATTTGCAACCTGCTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1509
 1460 CTTTCCCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1519
 1510 CTTTC-----TGAGGCTGCTTACTTACCAGG---ACTACAGGAGGCTGTGGAGGAGCTGTG 1561
 1520 GGAAGCAGAGCCCTAAGGATTAAAGCTCAGACACTCT 1558
 1562 AGAAATAGAGCCCTAAGGAGCTTACGCTCAGACACTCT 1600

RESULT 11
 AX381072 527 bp DNA Linear PAT 18-MAR-2002
 LOCUS AX381072
 DEFINITION Sequence 10 from Patent WO0212280.
 ACCESSION AX381072
 VERSION AX381072.1 GI:19575891
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 Pyle, R.A., Xu, J. and Secrist, H.
 AUTHORS
 TITLES
 JOURNAL
 PATENT: WO 0212280-A 10 14-FEB-2002;
 CORIXA CORPORATION (US)
 FEATURES
 source
 1..527
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 86 a 190 c 158 g 88 t 5 others
 ORIGIN

QY 397 CCGCTCCCGCCAGGACCATATGCTCAGCTTAATGTAGCCGATCCGGAGACAGTTT 456
 Db 348 CCGCTGTCCCGCCAGGAGGATTTGTCTTTCGGTATGACACCCGACCAAGACAGTTT 407
 QY 457 GTGCAATATGTGCGAATATCTCTACAGAGAGCTGGAATCTACCTGACCATCTGCGAG 516
 Db 408 GCAAGATTTCCCGCAATATCTCTAATATGACATGGAACCATCTCTCCACTGCGCAG 467
 QY 517 TGTGCGCCCTGTGACCCAGTATGGGCGCTGAGAGAGATGGCCCTGACAGCAAGCA 576
 Db 468 TGTGCGCCCTGTGACATTTGTCTGTGAGAGAGTTGCGCTTCCACAGGATC 527
 QY 577 GGAAGACCAAGTCCGCTGCGACCGGGAATTTCTGTGCTGCTGCGCTCGAGTTGA 636
 Db 528 GGAAGACCAAGTCCGCTGCGACCGGGAATTTCTGTGATCTGACATGAGTTG 587
 QY 637 CACACTG-----CGAGTACTTTGACTGCGCCGCTGGCACTGAAGCCGAGTCAAG 690
 Db 588 TGTACTGTAGAGAGAGCGGCTTGTACTGTCCAGGCTGGCAAGAGCGAGTCCAG 647
 QY 691 ATGAAGTTGGAGAGGATTAACAACACTGCGTCCCTGCAAGGAGGCGACTTCCAGAA 750
 Db 648 ATGAATTTATGATTAAGTACTGACGTCACACTGTGCTGTAAGCGGAGCACTTCCAG 707
 QY 751 CCGTCCCGCCCGAGCGCGCTGCGACCCGACACAGAGTGTGAGAACAGGTTGCTG 810
 Db 708 CTTCCTCCCGCTGAGCGCGCTGTCACACCCATACAGATGAGATTCAGGCGCTG 767
 QY 811 AGGAGCTCCAGGACACTGCGCTGCGACACAACTGCAAAATTCATTAGAGCACTG 870
 Db 768 AGGAGCTCCAGGATCTCTCTGATGATCATCTGTAAATCCCGAGAGCCAGAG 827
 QY 871 CCCCAGATGTGACAGATCGCTCAAG----- 899
 Db 828 CATGCTACTGCTGCGCACTCTGTGCTGCTCTCTTCTGCTTCAACACTGTC 887
 QY 900 AGGCTCCGAGGAGGAGAGGAGCCCAATCTGATGAGAGCTGGAGC----- 948
 Db 888 TGGCTGTGCTGAGTGAAGGACCCGCTCTCTGAGAAACCTGGTACTGCTGCAAGC 947
 QY 949 -----CTCGAAGGCCCATCATACTTCC 972
 Db 948 GGCACCCAGAGGAGAAATCTCCCTTCCCGCTCCCAAGAGCGGACCAATTTCC 1007
 QY 973 CTGACTTGTACAGCACTGCTACCATTTCTGAGAGATTTTCCCAAGTATCACTGAGC 1032
 Db 1008 CTGACTGTGAGAGCACTCTACCATGCTGTGAGACTTGTCCCAATCCCTGCTGAC 1067
 QY 1033 TCCCGCAGCGCCAGTTTGTGAGGAGGCGCGCAACAGACAGTCTCTGAGCTGA 1092
 Db 1068 CCCCAGCGCCCTCTCTGAGAGAGGAGTGTCTACACAGCAAGTCCCTGCTCCAGG 1127
 QY 1093 CCAAGAGCGCGAGTTGGAACCCGAGGAGAGAGAGTGGCCAGGATACCAAGGCA 1152
 Db 1128 CCAAGAGCTGTGAGCTGAGACTGTGAGAAATGAGCGAGGCGGCAAGGTCGAAATGCA 1187
 QY 1153 TTGATGTACCGCGGCTGATGACTTCACTGCAACATCTACATCTACATGAGAGCAG 1212
 Db 1188 TTGATGTACCGCGGCTGTGACTGTCTACCGCAACATCTACATCTACATGAGAGCAG 1247
 QY 1213 TACTGAGGAGGAGACCGGCTGTGAGACTCCAGGATCCCGGAAACCTCCATACCCCA 1272
 Db 1248 TGTGAGGAGGAGACCGGCGCTGTGAGAGCCCTGAGAGCTCCATACCCCA 1307
 QY 1273 TTCCCGAAGAGGAGAGCTGCGCTCCGCGGCTCTCTACACCCCAACAGAGATGAGA 1332
 Db 1308 CTCTCTGAAGAGGAGCTCTGCGCTCTGTGAGCTGTCTACACCTACAGAGAGATGAGA 1367
 QY 1333 AGGCTTGGCACTAGCGGAGAGAGAG--CACTGTGTGTGCAACACCTCTACAGAGGAGGCA 1391
 Db 1368 AAGCTTGGCACTTGTGCGCGAGAGAGCACTAGAGTGCAGAAAGCTCTGACAGAGG--CCA 1425

QY 1392 AGGAACCAATTTATACCCATGATGACTGAGTGTGAGAAAAGGACAGAGAGGGGCA 1451
 Db 1426 AGGAACCAATTTGACACCGCTGCTGAGGGCACTGGAAAAGGACAGAGAGGCTCC 1485
 QY 1452 CAAGGAGCACTTCTCCCTTGGAGGCGCTGCGCACGCTGGGATTTACAGGGGCTGAGTA 1511
 Db 1486 AGCAGCTTCTTTC-----TGAGGCTGTCTACTGACCAAG--ACTCACAGGCTGTGGG 1537
 QY 1512 GGGCCCGGGGAGAGAGCCCTTAAGGATTAAGGCTGACAGACCTT--GAGAGAGGT 1568
 Db 1538 AGGCTGTGAGAAATGAGAGCCCTTAAGGAGTATAGCTTACAGACCTCTTACAGAGAAC 1597
 QY 1569 GGGCACTGGCTGGGATGAGGCTGCTCCACAGAGACTTCTCTAGTCCCTGAGCAAACTGA 1628
 Db 1598 GGCACACTGACAGGCGCACAGT--CTGCGCCAGAGTCTGACTACCGCTGAGCAAGCCTGA 1656
 QY 1629 GGCCTCCCGGAGAGACCCACCCCGGGGGCTGCTACGCTTACGACGAGAGAGGCA 1688
 Db 1657 GACCTAGTTTACAGACCCGCTTCCAGAGGCTTGGCCATTTGAGCTGGACACAAAG 1716
 QY 1689 CATGATACCAACTGCTGCGCACTACAGACGCGCACCGGAGACAGGAGAGAGC 1747
 Db 1717 AAGGCTGTGACCCCGCG--CTGCTCTCTGTGTCTTAAACATGACAGAGAGAGCT 1774
 QY 1748 CGCCACAGGTCACCTGCAAGAGAGTACAGGAGGCGCTCTTAAGGATTTGTGTCTCATC 1807
 Db 1775 GGTCCCAAGGCTACTGCAAGAGATATCCAGGAGACCTGCAAAAGCTTTGTGTGTCTAT 1833
 QY 1808 CCCAAGCTTACAGAGACCTTTGGGGTTCCACACTTCACTGAGTGGAGTGAAGACCTGCA 1867
 Db 1834 CCTACTTACAGAGCGCTTGAAG--CTCATATTCTGCTGAGCAAGATATGACCTGGG 1889
 QY 1868 TGAAGATGAATTTATGAGAGAGAGCTCTCTCCCTCCCTCTTGAAGAGAGAGAGGA 1927
 Db 1890 GGAAGGTGAAGTGTGCGGAGAGAGCAACCCCTCTCTGCTGAGGAGAGATGAGA 1949
 QY 1928 GTCAATTAACAAGTGGGGGTTGGGTAGGATCTCTAGTATGGGAGAGAGTTGG 1962
 Db 1950 GAAGGAGAGA--TGGGGCTTGGGTAGCGTTAGCATGTGTGAGAGAGTTGG 2003

RESULT 10

MUSLYMPHOB

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PubMed

FEATURES

source

CDS

Location/Qualifiers
 1..1614
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /cell_line="Str-2"
 186..1433
 /note="putative"
 /codon_start=1
 /product="Lymphotoxin-beta receptor"
 /protein_id="AA080846.1"
 /db_xref="GI:600223"
 /translation="MLRPAASSPFGILAMGFLILGLSLIVASOPOLVPPYRIENQTCM
 DQKEVEYPMHIVCCSRCPGGEVFVAVCSRSODTVCKTQPHNSVNEHMHLSITQGLCR

```

STS      20971..21045
          /standard.name="D12S111E"
STS      22216..22330
          /standard.name="Cda1e01"
STS      22518..22670
          /standard.name="STS4862"
          complement(join(22595..22710,23806..23889,25797..26099,
          27000..27338,31028..31297,31447..31677,32432..32495))
gene
Query Match      42.5%; Score 841.4; DB 9; Length 140026;
Best Local Similarity 99.9%; Pred. No. 3.7e-160;
Matches 842; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1140 GGTACCAATGCGATTCATTCACCGCGGGCTATGACTATACATGCAACATCTACATC 1199
DB 94085 GGTACCAATGCGATTCATTCACCGCGGGCTATGACTATACATGCAACATCTACATC 94026

QY 1200 TACATGACACAGTACTGCGGGGACACCGGGTCTGTGAGACCTCCACCTACCCCGAA 1259
DB 94025 TACATGACACAGTACTGCGGGGACACCGGGTCTGTGAGACCTCCACCTACCCCGAA 93966

QY 1260 CCTCCATACCCCATTCGCCAAGAGGGGGACCTGGCCCTCCCGGGCTCTCTACACCCAC 1319
DB 93965 CCTCCATACCCCATTCGCCAAGAGGGGGACCTGGCCCTCCCGGGCTCTCTACACCCAC 93906

QY 1320 CAGGAAGATGGCAGAGCTTGCGACCTAGCGGAGACAGACATGTCACACACCTCT 1379
DB 93905 CAGGAAGATGGCAGAGCTTGCGACCTAGCGGAGACAGACATGTCACACACCTCT 93846

QY 1380 AACAGGGGCCCAAGAGCAACATTTATCACCATGACTGACTGAGTCTGAGAAAAAGGACA 1439
DB 93845 AACAGGGGCCCAAGAGCAACATTTATCACCATGACTGACTGAGTCTGAGAAAAAGGACA 93786

QY 1440 AGAAGGGGGGCAAGAGGACCTTCCTCTGAGGTCGCGCCCTCCACCTGGGATTCACA 1499
DB 93785 AGAAGGGGGGCAAGAGGACCTTCCTCTGAGGTCGCGCCCTCCACCTGGGATTCACA 93726

QY 1500 GGGGCTGTAGTAGGGCCCGGGAGAGAGCCCTAAGGATTAAAGCTAGACACCTCTG 1559
DB 93725 GGGGCTGTAGTAGGGCCCGGGAGAGAGCCCTAAGGATTAAAGCTAGACACCTCTG 93666

QY 1560 AAGACAGGTGGGCACTGGCTGGGTAGGTCCTCCACAGACTCTCCTACTGCTGAG 1619
DB 93665 AAGACAGGTGGGCACTGGCTGGGTAGGTCCTCCACAGACTCTCCTACTGCTGAG 93606

QY 1620 CAACCTGAGGCTCCCGGACAGACCCACCCCTGGGGTGTCACTCAGCTCAGGCACG 1679
DB 93605 CAACCTGAGGCTCCCGGACAGACCCACCCCTGGGGTGTCACTCAGCTCAGGCACG 93546

QY 1680 GACAGGGACATGATACCAACTGTGCCCCACTACAGACAGCCGACCGGAGCAGGCAC 1739
DB 93545 GACAGGGACATGATACCAACTGTGCCCCACTACAGGACAGCCGACCGGAGCAGGCAC 93486

QY 1740 GAGGAGGCGCCACACAGGTTCACCTGCAAGAGCTCAGCGGCCCTCTTAAAGATTCTGG 1799
DB 93485 GAGGAGGCGCCACACAGGTTCACCTGCAAGAGCTCAGCGGCCCTCTTAAAGATTCTGG 93426

QY 1800 TGTCTATCCCAAGCTTGAGAGCCCTTGGGGTTCACACTTCACTGAGTGTGAGGTGAG 1859
DB 93425 TGTCTATCCCAAGCTTGAGAGCCCTTGGGGTTCACACTTCACTGAGTGTGAGGTGAG 93366

QY 1860 ACCCTGACATGAAGATGAATTTATAGGAGAGAGCTCTCCCTCCCTCTAGAGAGAG 1919
DB 93365 ACCCTGACATGAAGATGAATTTATAGGAGAGAGAGCTCTCCCTCCCTCTAGAGAGAG 93306

QY 1920 GAAAGGAGATCTTAAACAACACTAGGGGGTGGGTAGGATTCCTAGATGAGGAGAGTTT 1979
DB 93305 GAAAGGAGATCTTAAACAACACTAGGGGGTGGGTAGGATTCCTAGATGAGGAGAGTTT 93246

QY 1980 TGG 1982
DB 93245 TGG 93243

```

```

RESULT 9
LOCUS   MMU29173
DEFINITION Mus musculus lymphotoxin-beta receptor mRNA, complete cds.
ACCESSION U29173
VERSION 1
KEYWORDS GI:881620
SOURCE  Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2076)
AUTHORS  Force, W.R., Williams-Abbott, L., Browning, J., Hession, C., Tizard, R.
          and Ware, C.F.
TITLE    Cloning and characterization of the Mouse lymphotoxin-beta Receptor
JOURNAL  Unpublished
REFERENCE 2 (bases 1 to 2076)
AUTHORS  Force, W.R.
TITLE    Direct Submission
JOURNAL  Submitted (15-JUN-1995) Walker R. Force, Biomedical Sciences, U.C.
          Riverside, Riverside, CA 92521-0121, USA
FEATURES
         source
            1..2076
            /organism="Mus musculus"
            /strain="CVB"
            /db_xref="taxon:10090"
            /sex="female"
            /tissue_type="lung"
            /note="transmembrane protein"
            /product="lymphotoxin-beta receptor"
            /protein_id="AA06964.1"
            /codon_start=1
            /translation="MRLPRASSPCGLANGPILLGLISGLIVASQPOLVPPRIENQTCM
            /db_xref="GI:881621"
            /translation="MRLPRASSPCGLANGPILLGLISGLIVASQPOLVPPRIENQTCM
            DQREYEPMDVCCSRCPGEFVAVCSRDQIVYCKTSPHNSYNENHNLSTQLCR
            PCDIYLFGEFEVAPCTSDRKACRCOPGSCYLDNECVHCEERLYLQPTAEVTD
            EIMPTDVNCVPCRGHFOGTSRPARCOPHRCFOGLVEAPGYSYDCTCKNPEP
            GAMILLATILSLVLETLFTVLACAMHNPGLCKRLGLTLRHPEGEGSPCPAPRAD
            PHFDIAPLIPMSGDLSPPAGPPTAPSLSEVYLQOOSPIVQARELEAPGEGHOVA
            HGANGIHVTSVTVTGNIITYNGFVLGTPGDPAPPEPPTPEGAPRSELS
            TPYQEDGKAMHLAETLGLGCDL"
BASE COUNT  436 a      652 c      597 g      391 t
ORIGIN
Query Match      38.6%; Score 765.8; DB 10; Length 2076;
Best Local Similarity 67.8%; Pred. No. 1.1e-144;
Matches 1379; Conservative 0; Mismatches 532; Indels 124; Gaps 17;

QY 40 CCGCATGGAGAGCCCGTGGAGGCGCGGCTGCCGCGCTGGGGTGCACATGGGC 99
DB 1 CCGGCGCTGAGACCTCTGGAGGCGCAGGCGCGGCTGCCGCGCTGGGGTGCACATGGGC 59

QY 100 CCTGAGTCCGTCAGAGGCTCTGGGCTGGGAGCGCCGCGCACGCTGCGCAGGAGCT- 158
DB 60 -----CCAGCGCGCAGGCTTGGAAGTCGTGGCCACACGCTGTCAGGACGTC 107

QY 159 CGGGCTCTGCTGCTCTCTCCAGGCCCA--CGTTGCTGGCGGCTGGCCGAGTGGCGG 216
DB 108 CGGGCTCTGCTGCTCTCTCTCTAGACTCCACCGGCTGTGAGCTCAGCGACCCGAAAGGCGG 167

QY 217 CCATGCTCTGCTGCTGGGACACCTGCGCCCGGGCTGGGAGGCTGTGGTGG 276
DB 168 CCATGCGCTGCTGCTGCGGCGCTCTCTCTGCGGCTGGCTGGGAGGCTACTCTGCTGG 227

QY 277 GCGTCTTGGGGCTCTGTGAGAGATGCGACCCCGAGGCGTGGCTCCATATGCGTGAGA 336
DB 228 GACTCAGCGGGCTCTGTGGGCTCTCAGCCCGAGGTGGTGGCCCTTATGCAATGAAA 287

QY 337 ACCAGACCTGACGAGCAGCAAGAGATTAATAGAGCCCGACAGCCGACATCTGCTGT 396
DB 288 ACCAGACCTGCTGGGACCGAGCAAGGATTAATAGAGCCCGACAGCGAGTCTGCTGT 347

```

across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Attempts are made to complete double stranded sequence for all regions. All sequence is completed to a standard of coverage with a minimum of 3 reads with no ambiguities. If the sequence coverage for a region does not meet this standard, it is indicated in the annotation as Low Coverage. Low coverage linkages are verified by PCR product size verification or verification of forward and reverse reads from clones which span the low coverage area.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated average error rate is less than 1 per 10,000 bases using the Consed quality parameters. Regions that do not meet this requirement are annotated as Low Quality.

-----Summary Statistics

Center project name: RPI-102E24
Sequencing vector: M13
Chemistry: Dye-terminator Big Dye: 100%
Assembly program: Phrap version 0.990319
Contig length: 140026
Fraction of Phrap value < 40: 0.0426
Error Rate in Consed: 0.56 per 10,000 bases
Number of N's in consensus: 1

----- Distribution of Quality < 40 Bases:

	5	10	15	20	25	30	35	40
1000							*	*
900							*	*
800							*	*
700							*	*
600							*	*
500							*	*
400					*	*	*	*
300				*	*	*	*	*
200			*	*	*	*	*	*
100		*	*	*	*	*	*	*
0	*	*	*	*	*	*	*	*

FEATURES

Source Location/Qualifiers
1. 140026
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RPI-102E24"
1. 51
/rpt_family="AluJb"
52..76
/rpt_family="(CAAA)n"
complement(254..367)
/rpt_family="LIMB8"
complement(377..649)
/rpt_family="LIMB8"
complement(650..941)
/rpt_family="AluSc"
complement(983..1090)
/rpt_family="LIMB8"
complement(1110..1339)
/rpt_family="AluSc"
complement(1341..1354)
/rpt_family="LIMB8"
complement(1355..1652)
/rpt_family="AluSc"
complement(1653..1787)
/rpt_family="LIMB8"
complement(1797..2090)
/rpt_family="AluJc"
complement(2241..2533)

repeat_region /rpt_family="AluSg"
3376..3664
repeat_region /rpt_family="AluSg"
3825..3850
gene /rpt_family="AT-rich"
complement(join(4534..4884,4933..5259,5275..5502,
5517..5666,5694..5849,5883..6026,6036..6368,6473..6883,
6889..7029))
/gene="Plakophilin 2 (PKP2)"
7134..7432
repeat_region /rpt_family="AluSc"
complement(7442..7588)
repeat_region /rpt_family="LIM4"
8060..8245
repeat_region /rpt_family="LIMed"
8246..8360
repeat_region /rpt_family="HAL1"
8399..8563
repeat_region /rpt_family="AluSg"
8564..9933
repeat_region /rpt_family="SVA"
9934..10088
repeat_region /rpt_family="LIMD2"
10107..10188
repeat_region /rpt_family="FLAM_A"
10194..10975
repeat_region /rpt_family="LIMD2"
10976..11283
repeat_region /rpt_family="AluY"
complement(11303..11589)
repeat_region /rpt_family="AluJb"
complement(11678..11692)
repeat_region /rpt_family="FRAM"
complement(11693..11995)
repeat_region /rpt_family="AluSc"
complement(11996..12157)
repeat_region /rpt_family="FRAM"
12547..12899
repeat_region /rpt_family="MER104"
complement(12900..13198)
repeat_region /rpt_family="AluSc"
13199..13254
repeat_region /rpt_family="MER104"
complement(13309..13472)
repeat_region /rpt_family="L2"
13793..14017
repeat_region /standard_name="GDB:384731"
complement(16498..17181)
repeat_region /rpt_family="THE1A-int"
complement(17193..17563)
repeat_region /rpt_family="THE1A-int"
complement(17564..17915)
repeat_region /rpt_family="THE1A"
complement(18210..18342)
repeat_region /rpt_family="FLAM_A"
join(18391..18519,18744..18902,19803..19853)
/gene="Vesicle-associated membrane protein 1 (SVB1)"
19248..19275
repeat_region /rpt_family="(CA)n"
19276..19351
repeat_region /rpt_family="LIMed"
complement(19296..19475)
repeat_region /rpt_family="LIMCa"
19514..19933
repeat_region /rpt_family="(CCG)n"
20287..20651
repeat_region /standard_name="GDB:214816"
20290..20527
repeat_region /standard_name="sts-M36200"
20459..20650
repeat_region /standard_name="sts-M36200"
20841..21018
repeat_region /standard_name="SHC-557"

QY 680 CGAGCTCAAGATGAAAGTTGGGAAGGTAACAACCACTGCGTCCCTGCAAGGAGGCA 739
 |||||
 Db 617 CGAGCTCAAGATGAAAGTTGGGAAGGTAACAACCACTGCGTCCCTGCAAGGAGGCA 676
 |||||
 QY 740 CTTCAGAAATACCTCTCCCGAGCGCGCTGCCAGCCCAACACAGGTGTGAGAACCA 799
 |||||
 Db 677 CTTCAGAAATACCTCTCCCGAGCGCGCTGCCAGCCCAACACAGGTGTGAGAACCA 736
 |||||
 QY 800 AGGTGGTGGAGGAGGAGCTCCAGGAGTGCAGTCCGACACAACTGCAAAATCAAT 859
 |||||
 Db 737 AGGTGGTGGAGGAGGAGCTCCAGGAGTGCAGTCCGACACAACTGCAAAATCAAT 796
 |||||
 QY 860 AGAGCACTGCCCCAGAGATGTCA----- 884
 |||||
 Db 797 AGAGCACTGCCCCAGAGATGTCA----- 856
 |||||
 QY 885 ----- 884
 |||||
 Db 857 GGCTTTCTTCTGCTCTTGCACCGTCTTCTCTGATCTGGAAGAGCCACTTCTCT 916
 |||||
 QY 885 -----GGATGCTGCTCAAGAGGCTCCGAGGAGGAGGAGCCCAATCTGT 931
 |||||
 Db 917 CTGCAAGAACTGGATGCTGCTCAAGAGGCTCCGAGGAGGAGGAGCCCAATCTGT 976
 |||||
 QY 932 AGCTGGAAGCTGGAGCTCCGAAAGCCCAATCATCTTCTCTGATCTGTACAGCCACT 991
 |||||
 Db 977 AGCTGGAAGCTGGAGCTCCGAAAGCCCAATCATCTTCTCTGATCTGTACAGCCACT 1036
 |||||
 QY 992 GCTACCACTTCTGAGATGTTTCCCAAGTATCACTGGGCTCCCGAGCCCAAGTTT 1051
 |||||
 Db 1037 GCTACCACTTCTGAGATGTTTCCCAAGTATCACTGGGCTCCCGAGCCCAAGTTT 1096
 |||||
 QY 1052 GGAGCAGAGGCTGGCCGAACAGCAGATCTGTGACCTGACAGGAGGAGCCGACATGGA 1111
 |||||
 Db 1097 GGAGCAGAGGCTGGCCGAACAGCAGATCTGTGACCTGACAGGAGGAGCCGACATGGA 1156
 |||||
 QY 1112 ACCCGGAGAGCAGAGCAGGTTGGCCCAAGTATCACTTCTGTACAGGAGGAGTTC 1171
 |||||
 Db 1157 ACCCGGAGAGCAGAGCAGGTTGGCCCAAGTATCACTTCTGTACAGGAGGAGTTC 1216
 |||||
 QY 1172 TATGACTATCACTGCAACATCTATCATATGATGAGCAGTGTGGGAGCACCAGG 1231
 |||||
 Db 1217 TATGACTATCACTGCAACATCTATCATATGATGAGCAGTGTGGGAGCACCAGG 1276
 |||||
 QY 1232 TCCTGAGAGCTCCAGCTACCCCGGAACCTCATACCCCAATTCGGAAGAGGAGGAGCC 1291
 |||||
 Db 1277 TCCTGAGAGCTCCAGCTACCCCGGAACCTCATACCCCAATTCGGAAGAGGAGGAGCC 1336
 |||||
 QY 1292 TGGCCCTCCCGGCTCTTACACCCCAAGAGATGAGCAGGCTTGGCAGCTAGCGGA 1351
 |||||
 Db 1337 TGGCCCTCCCGGCTCTTACACCCCAAGAGATGAGCAGGCTTGGCAGCTAGCGGA 1396
 |||||
 QY 1352 GACAGAGCACTGTGTGTCACACCCCTTACACAGGAGGAGGAGGAGGAGGAGGAGGAGG 1411
 |||||
 Db 1397 GACAGAGCACTGTGTGTCACACCCCTTACACAGGAGGAGGAGGAGGAGGAGGAGGAGG 1456
 |||||
 QY 1412 TGACTGACTGAGTGTGAGAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1471
 |||||
 Db 1457 TGACTGAGGAGTGTGAGAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1516
 |||||
 QY 1472 AGGCTGCTGTCACAGTGGAGTTCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1531
 |||||
 Db 1517 AGGCTGCTGTCACAGTGGAGTTCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1576
 |||||
 QY 1532 CTAAGGAGTTTAAAGCTCAGACACCTCAGAGAGTGGGCTGTGGTGTAGCGTGGC 1591
 |||||
 Db 1577 CTAAGGAGTTTAAAGCTCAGACACCTCAGAGAGTGGGCTGTGGTGTAGCGTGGC 1636
 |||||
 QY 1592 CTCCACAGAGACTCTCCCTACTGCTGAGCAAACTGAGGAGCTCCCGGAGAGCCCAAGC 1651
 |||||
 Db 1637 CTCCACAGAGACTCTCCCTACTGCTGAGCAAACTGAGGAGCTCCCGGAGAGCCCAAGC 1696
 |||||

QY 1552 CCCCTGGGGCTGCTCAGCCTCAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1711
 |||||
 Db 1697 CCCCTGGGGCTGCTCAGCCTCAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1756
 |||||
 QY 1712 ACAGACGCGCGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1771
 |||||
 Db 1757 ACAGACGCGCGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1816
 |||||
 QY 1772 GTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1831
 |||||
 Db 1817 GTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1876
 |||||
 QY 1832 GTTCAGACCTTCAGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1891
 |||||
 Db 1877 GTTCAGACCTTCAGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1936
 |||||
 QY 1892 GCTCTTCCCTCCCTCCTCTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1951
 |||||
 Db 1937 GCTCTTCCCTCCCTCCTCTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1996
 |||||
 QY 1952 TAGATCTCAGTATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1982
 |||||
 Db 1997 TAGATCTCAGTATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2027
 |||||

RESULT 8

AC005840/c

LOCUS

DEFINITION

SYB1, CD27, and SCNN1A genes.

AC005840

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

ANNOTATION OF FEATURES:

STS are identified using ePCR (genome Res. 7:541-550).

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for human sequences.

Genes and Regions of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to EST and cDNA sequences in Unigene. Genes demonstrate at least two exons flanked by consensus splice sites that maintain sequence continuity

DB	Accession	Species	Gene	Length (bp)	Insertion (bp)	Position (bp)	Sequence	Reference
OY	1263	CCATACCCCATTC	CCCAAGAGGGGGAGCC	1322			CGGCGCCCGGGGCTCTACACCCACACG	1322
Db	1321	CCATACCCCATTC	CCCAAGAGGGGGAGCC	1380			CGGCGCCCGGGGCTCTACACCCACACG	1380
OY	1323	GAAGATGCGAAGG	CTTGGCACCCTAGCGG	1382			GACAGACACTGTGTGTCACACCCCTTAAC	1382
Db	1381	GAAGATGCGAAGG	CTTGGCACCCTAGCGG	1440			GACAGACACTGTGTGTCACACCCCTTAAC	1440
OY	1383	AGGGGGCCCAAG	GAACCAATTATACCC	1442			TGACTGACTGTGTGAGAAAAGGCAGAA	1442
Db	1441	AGGGGGCCCAAG	GAACCAATTATACCC	1500			TGACTGACTGTGTGAGAAAAGGCAGAA	1500
OY	1443	AGGGGGCCCAAG	GGGACACCTTCTCC	1502			CTTCAAGGCTGGCCCTGACACCTGGATTCACAGG	1502
Db	1501	AGGGGGCCCAAG	GGGACACCTTCTCC	1560			CTTCAAGGCTGGCCCTGACACCTGGATTCACAGG	1560
OY	1503	GCCTAGTAGTGG	GGCCGGGGAAGCAG	1562			AGCCCTTAAGGATTAAGGCTCAGACACCTTGAGA	1562
Db	1561	GCCTAGTAGTGG	GGCCGGGGAAGCAG	1620			AGCCCTTAAGGATTAAGGCTCAGACACCTTGAGA	1620
OY	1563	GCAGGTGGGCAT	GTGCTGGGTACGG	1622			CGCCCTCACAGACTCCCTACTGCTGTGACAA	1622
Db	1621	GCAGGTGGGCAT	GTGCTGGGTACGG	1680			CGCCCTCACAGACTCCCTACTGCTGTGACAA	1680
OY	1623	ACCTAGAGCCTC	CCGGCAGACCCACC	1682			CCCTGGGCTGTCAAGCCTCAGAGCAAGGAC	1682
Db	1681	ACCTAGAGCCTC	CCGGCAGACCCACC	1740			CCCTGGGCTGTGTCAAGCCTCAGAGCAAGGAC	1740
OY	1683	AGGGGACATGAT	AACACGCTGGCC	1742			CACTACAGCACCGGCACCGGAGACGGCACCGAG	1742
Db	1741	AGGGGACATGAT	AACACGCTGGCC	1800			CACTACAGCACCGGCACCGGAGACGGCACCGAG	1800
OY	1743	GGAGCGCGCCAC	AGGTCACCTCCTCA	1802			AGGAGCGTCACGGGCCCTCTTAAGGATTCGTGTGC	1802
Db	1801	GGAGCGCGCCAC	AGGTCACCTCCTCA	1860			AGGAGCGTCACGGGCCCTCTTAAGGATTCGTGTGC	1860
OY	1803	TCATCCCCAAGT	TAGAGACCTTTGG	1862			GTCCACACTTACAGTGTGAGCTGAGGTAGACC	1862
Db	1861	TCATCCCCAAGT	TAGAGACCTTTGG	1920			GTCCACACTTACAGTGTGAGCTGAGGTAGACC	1920
OY	1863	CTGCATGGAAG	ATAATATAGGAG	1922			AGCGCTCTCCCTCCCTCTCTAGAGAGAGGAA	1922
Db	1921	CTGCATGGAAG	ATAATATAGGAG	1980			AGCGCTCTCCCTCCCTCTCTAGAGAGAGGAA	1980
OY	1923	AGGGGATCATTA	TAACTAGGAGG	1982			GTGGTGGGTAGGATTCCTAGGTATGGGAGAGCTTTGG	1982
Db	1981	AGGGGATCATTA	TAACTAGGAGG	2040			GTGGTGGGTAGGATTCCTAGGTATGGGAGAGCTTTGG	2040

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
2 (bases 1 to 2091)	Isogai,T. and Yamamoto,J.	Direct Submission	Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team) ; 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)	
			NEO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan: CDNA full insert sequencing: Research Association for Biotechnology (RAB) ; cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'-, 6'-, 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.	
FEATURES				
SOURCE				
				1..2091
				/organism="Homo sapiens"
				/db_xref="taxon:9606"
				/clone="D30S71000238"
				/cell_type="CD34+ Cells"
				/clone_id="D30S71"
				/note="cloning vector: pME18SFL3-mRNA from CD34+ cells after 3-days ODF induction, primary culture, CD34+ Cells"
BASE COUNT				440 a 691 c 594 g 366 t
ORIGIN				
Query Match				87.4%; Score 1732.2; DB 9; Length 2091;
Best Local Similarity				92.5%; Pred. No. 0;
Matches 1916; Conservative				0; Mismatches 3; Indels 152; Gaps 3;
QY	20	CCACTCCCACTTCTCTGAGCTCCGCCATGAGAGCCCTGGAGGCCCGCTTGCCCTCCG	79	
Db	1	CCACTCCCACTTCTCTGAGCTCCGCCATGAGAGCCCTGGAGGCCCGCTTGCCCTCCG	60	
QY	80	GCCCTGGGGTGCATATCGGCGCCCTGAGTCCCGCTCCAGAGCTCTGGGGTCGGGACAGCGCG	139	
Db	61	GCCCTGGGGTGCATATCGGCGCCCTGAGTCCCGCTCCAGAGCTCTGGGGTCGGGACAGCGCG	120	
QY	140	CCACCGCTGCCAGAGAGCGTGGGGCTCTTGCTTCCGCCAGGCGCCACGTTGCTGGCC	199	
Db	121	CCACCGCTGCCAGAGAGCGTGGGGCTCTTGCTTCCGCCAGGCGCCACGTTGCTGGCC	180	
QY	200	GCTTGGCGGAGTGTGGCGGCATATGCTCTCGCTTGGGGCCACGCTGGCCCGGCGTGGCGT	259	
Db	181	GCTTGGCGGAGTGTGGCGGCATATGCTCTCGCTTGGGGCCACGCTGGCCCGGCGTGGCGT	212	
QY	260	GGGGCTCTGTGCTGTGGGCTCTTCTGGGCTCTTGAGACATCGAGGCCAGAGCGGTGCC	319	
Db	213	GGGGCTCTGTGCTGTGGGCTCTTCTGGGCTCTTGAGACATCGAGGCCAGAGCGGTGCC	271	
QY	320	TTCATATGCGTCGGGAGAAACAGACACTCTCAGGAGACAGAGAAAGAAATCTATGAGCCCA	379	
Db	272	TTCATATGCGTCGGGAGAAACAGACACTCTCAGGAGACAGAGAAAGAAATCTATGAGCCCA	331	
QY	380	GCACCGCATCTGCTGCTCCGCTGCCGCTGCCAGGACACTATGTCTCCAGCTTAATGTAGCG	439	
Db	332	GCACCGCATCTGCTGCTCCGCTGCCGCTGCCAGGACACTATGTCTCCAGCTTAATGTAGCG	376	
QY	440	CATTCGGGACACAGTTTGTGCCACATGTGCCGAGATTTCTTACACAGACACTGGAACTA	499	
Db	377	CATTCGGGACACAGTTTGTGCCACATGTGCCGAGATTTCTTACACAGACACTGGAACTA	436	
QY	500	CGTACCATCTGGCAGGTGTGGCGCCCGCTGTGACCCAGTATGGCGCTCAGAGAGATTGC	559	
Db	437	CGTACCATCTGGCAGGTGTGGCGCCCGCTGTGACCCAGTATGGCGCTCAGAGAGATTGC	496	
QY	560	CCCTGACACAGCAAAAGGAAGACCCAGTCCGCTGCACGCGGAGATGTTCTGTCTGC	619	
Db	497	CCCTGACACAGCAAAAGGAAGACCCAGTCCGCTGCACGCGGAGATGTTCTGTCTGC	556	
QY	620	CTGGGGCTCTGAGTGTACACACTTGCAAGTCTTTCTGACTGCCCGCGCTGGACTGAAAGC	679	
Db	557	CTGGGGCTCTGAGTGTACACACTTGCAAGTCTTTCTGACTGCCCGCGCTGGACTGAAAGC	616	

OY	1863	CTGATGTGAAGTAAATTAATGGAGACGACTCCTTCCTCCCTCCCTCTTAGAGAGAGAA	1922
Db	1921	TTTGCGTGAAGATGAATAATTATAGGAGAGACGCTCCTTCCTCCCTCTTAGAGAGAGAA	1980
OY	1923	AGGAGCATTAACAACAACAGAGGGGCTTGCGTAGATCTCTAGGTATGGGAGAGATTGG	1982
Db	1981	AGGAGCATTAACAACAACAGAGGGGCTTGCGTAGATCTCTAGGTATGGGAGAGATTGG	2040
 RESULT 6 HUMFNRRP HUMFNRRP 2136 bp mRNA linear PRI 03-AUG-1993			
DEFINITION Homo sapiens (clone CD18) tumor necrosis factor receptor 2 related protein mRNA, complete cds.			
ACCESSION L04270.1 GI:339761			
VERSION L04270.1			
KEYWORDS tumor necrosis factor receptor 2 related protein.			
SOURCE Homo sapiens (library: liver cDNA of P.M.) liver cDNA to mRNA.			
ORGANISM Homo sapiens			
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 2136)			
AUTHORS Baens,M., Chalfanet,M., Cassiman,J.J., van den Berghe,H. and Marynen,P.			
TITLE Construction and evaluation of a hncDNA library of human l2p			
JOURNAL	Genomics 16 (1), 214-218 (1993)		
MEDLINE	93252381		
PUBMED	8486360		
FEATURES			
source location/Qualifiers			
1..2136			
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/tissue_type="liver"			
/tissue_1lb="liver cDNA of P.M."			
169..1476			
/note="putative"			
/codon_start=1			
/product="Tumor necrosis factor receptor 2 related protein"			
/protein_id="AAA36757.1"			
/db_xref="GI:339762"			
/translation="MLPLPASAPGLAMGPLVLFLFGLIASOPAVPYVASENOTQR DOEEHYEPQRIICSCSPREPITYSAKSRIRDYVCATCAENSNEHNITITQLCR PCDVMGLEIEPTCKTRKTDCRQPDPGFAMALECHTECLSDCPGETAEIKDEVE GKGNNHCVPCKAAGHFONTSSPSARCOPTHRENOGLVEAAGTAOSDTCKNLEPEE PEMSGKMIMLAVILPLAFELIATVFCISIMKSHPSLKRLSGSLKRPOGSPNVAVG SWBPBKAPHPYPDLVQPLLPISGDVSPYSTGLPAAPYLEACVPOOGSLDLTRPEOLE PGEOOVAHGNGIHVHTGGSMTITGNINYIVNGVIGGPBGDDLPAPIPEPPYPIPEES DPGPGSLTPHQDEGCKAMHLAEFHCHCATPENRKPRNDFTIHD"			
BASE COUNT 446 a 706 c 608 g 376 t			
ORIGIN			
Query Match 91.3%; Score 1809.2; DB 9; Length 2136;			
Best Local Similarity 94.6%; Pred. No. 0;			
Matches 1929; Conservative 0; Mismatches 3; Indels 108; Gaps 1;			
OY	51	GCCCTGAGAGGCCCGGCTGAGCCGCTCCCGGCTTGAGGATGCACATCGGCGCTTCCCG	110
Db	1	GCCCTGAGAGGCCCGGCTGAGCCGCTCCCGGCTTGAGGATGCACATCGGCGCTTCCCG	60
OY	111	TCCGAGGCTGAGGGGTGGGGGACGCGCGCACGCGCTGCCAAGAGAGTGGGCTTCGCG	170
Db	61	TCCAGGCTGAGGGGTGGGGGACGCGCGCACGCGCTGCCAAGAGAGTGGGCTTCGCG	120
OY	171	CTTCTCCACAGGCCCCACAGTTGCTGAGCGCGCTGCGCAGTAGTGCCGCGCATGCTCGCT	230
Db	121	CTTCTCCACAGGCCCCACAGTTGCTGAGCGCGCTGCGCAGTAGTGCCGCGCATGCTCGCT	180
OY	231	TGGGCACTCTGCCCCGGGCTGGCTGAGGAGGCTTGAGTGTGGGCTCTTGGGCTC	290
Db	181	TGGGCACTCTGCCCCGGGCTGGCTGAGGAGGCTTGAGTGTGGGCTCTTGGGCTC	240

QY	291	CTGCGAGCATGCGAGGCCCGCCAGGGGGTGGCTCTCCATATGCGTCCGGAAACACAGCACTGCGAGG	350
Db	241	CTGGAGAGATGCGAGCCCCCGAGGGGGTGGCTCCATATGCGTCCGGAAACACAGCACTGCGAGG	300
QY	351	GACCCAGGAAAGGAATACCTATGAGCCCCAGACACCCGATCTGGTGGCTCCCGCGCCGCA	410
Db	301	GACCAAGGAAAGGAATACCTATGAGCCCCAGACACCCGATCTGGTGGCTCCCGCGCCGCA	360
QY	411	GGCACATATGTCTCAGCTAAATGTAGCCGATCCGGGACACAGTTGTGCCAATGTGCC	470
Db	361	GGCACATATGTCTCAGCTAAATGTAGCCGATCCGGGACACAGTTGTGCCAATGTGCC	420
QY	471	GAGATTCCTCAACAGGAGCACTGGAACTACGTAGCCATCTGGCACTGTGCCGCCCTCGT	530
Db	421	GAGATTCCTCAACAGGAGCACTGGAACTACGTAGCCATCTGGCACTGTGCCGCCCTCGT	480
QY	531	GACCCAGTATGGGGCCCTCGAGGAGATTGGCCCCCTGCACAAAGCAAGGAGACCCAGTGC	590
Db	481	GACCCAGTATGGGGCCCTCGAGGAGATTGGCCCCCTGCACAAAGCAAGGAGACCCAGTGC	540
QY	591	CGCTGCGAGCCGGGAATGTTCTGTGCTGCTGGGCCCTTGAATGTACACACTGCGAGCTA	650
Db	541	CGCTGCGAGCCGGGAATGTTCTGTGCTGCTGGGCCCTTGAATGTACACACTGCGAGCTA	600
QY	651	CTTTCTGACTGCCCCCGCTGGGCACGTGAAGCCGAGCTCAAGATGAAAGTTGGGAAGGGTATAC	710
Db	601	CTTTCTGACTGCCCCCGCTGGGCACGTGAAGCCGAGCTCAAGATGAAAGTTGGGAAGGGTATAC	660
QY	711	AACCACTGCGTCCCCCTCGAGGAGGAGCACTTTCAGAAATACCTCCCTCCCGCAGCCCGC	770
Db	661	AACCACTGCGTCCCCCTCGAGGAGGAGCACTTTCAGAAATACCTCCCTCCCGCAGCCCGC	720
QY	771	TGCCAGCCCCACACACAGGTGTGAGAACCAAGTCTGTGTGGAGGCACTCCAGCAGCTGCC	830
Db	721	TGCCAGCCCCACACACAGGTGTGAGAACCAAGTCTGTGTGGAGGCACTCCAGCAGCTGCC	780
QY	831	CAGTCCGACACAACTCTGCAAAAATTCATTAGACCACTGCCCGCCAGAGATGCA-----	884
Db	781	CAGTCCGACACAACTCTGCAAAAATTCATTAGACCACTGCCCGCCAGAGATGCAAGAAC	840
QY	885	-----	884
Db	841	ATGCGATGCGTGGCGGCTTGTGCTGCCACTGCGGCTTCTTCTGCTCTCTGCCACGCTGTC	900
QY	885	-----GATCGCTGCTCAAGAG	902
Db	901	TCTTCGATCTGGAGAGCCACCCCTTCTCTCTGCGAGAAACTGGGATCGCTGCTCAAGAG	960
QY	909	CGTCCGAGGGAGAGAGAACCCAAATCCGTATGCTGGAAGTGGGAGACCTCCGAAGGCCAT	962
Db	961	CGTCCGAGGGAGAGAGAACCCAAATCCGTATGCTGGAAGTGGGAGACCTCCGAAGGCCAT	1020
QY	963	CCATACTTCCCTGACTTGGTATAGACGCACTGCTACCCATTCTTGAGATGTTTCCCACTA	1022
Db	1021	CCATACTTCCCTGACTTGGTATAGACGCACTGCTACCCATTCTTGAGATGTTTCCCACTA	1080
QY	1023	TCCACTGGGCTCCCGCAGCCCAAGTTTTSAGAGCGAGGGGTGCCGACAGCAGAGTCTT	1082
Db	1081	TCCACTGGGCTCCCGCAGCCCAAGTTTTSAGAGCGAGGGGTGCCGACAGCAGAGTCTT	1140
QY	1083	CTGGACCTGACACAGGAGGCGCAGTTGGAAACCCGGGAGACAGAGCCAGTGGCCCCAGGT	1142
Db	1141	CTGGACCTGACACAGGAGGCGCAGTTGGAAACCCGGGAGACAGAGCCAGTGGCCCCAGGT	1200
QY	1143	ACCAATGGCATTCATGTTCACCGCGGGGTCTATGACTATCATCTGGCAGACATCTACATCTAC	1202
Db	1201	ACCAATGGCATTCATGTTCACCGGGGGTCTATGACTATCATCTGGCAGACATCTACATCTAC	1260
QY	1203	AATGACCACTACTGTTGGGGGAGACACCGGGTCTCTGGAGACCTCCACGCTAACCCCGAAGCT	1262
Db	1261	AATGACCACTACTGTTGGGGGAGACACCGGGTCTCTGGAGACCTCCACGCTAACCCCGAAGCT	1320

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
1 Alvarés, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.
AUTHORS
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2135 11-APR-2002;
GENE LOGIC INC (US)

FEATURES
Location/Qualifiers
1. 2136

/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 446 a 706 c 608 g 376 t
ORIGIN

Query Match Best Local Similarity 94.6%; Pred. No. 0;
Matches 1929; Conservative 0; Mismatches 3; Indels 108; Gaps 1;

51 GCCCTGAGAGGCCCGCTGCGCTCCGCGCCCTGGGGGTCACATCGGCGCTGAGTCCCG 110
1 GCCCTGAGAGGCCCGCTGCGCTCCGCGCCCTGGGGGTCACATCGGCGCTGAGTCCCG 60
111 TCCGAGGCTGTGGGCTGGGCGACGCCGCCACCGCTGCCAGAGAGTGGGCTCTGTC 170
61 TCCGAGGCTGTGGGCTGGGCGACGCCGCCACCGCTGCCAGAGAGTGGGCTCTGTC 120
171 CTTCCTCCAGAGGCCCGCTGCTGGCGCGCTGGCGAGAGGCGCGCATCTGCTGCT 230
121 CTTCCTCCAGAGGCCCGCTGCTGGCGCGCTGGCGAGAGGCGCGCATCTGCTGCT 180
231 TGGGCGACCTCTGCGCCCGCGCTGGCGTGGGGGCTGTGTGTGTGGGCTCTTGGGCTC 290
181 TGGGCGACCTCTGCGCCCGCGCTGGCGTGGGGGCTGTGTGTGTGGGCTCTTGGGCTC 240
291 CTGGCAGCATGCGAGCGCCCGCGCTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGG 350
241 CTGGCAGCATGCGAGCGCCCGCGCTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGG 300
351 GACCAAGAAAGAAATATATATAGAGCCCGACCGCATCTGCTCTCCGCTGGCGCGCA 410
301 GACCAAGAAAGAAATATATATAGAGCCCGACCGCATCTGCTCTCCGCTGGCGCGCA 360
411 GGCACATATGTCACGTAATATATAGAGCCCGACCGCATCTGCTCTCCGCTGGCGCGCA 470
361 GGCACATATGTCACGTAATATATAGAGCCCGACCGCATCTGCTCTCCGCTGGCGCGCA 420
471 GAGAATTCCTACAGAGCAGTGAACCTGACATCTGACATCTGACATCTGACATCTGACAT 530
421 GAGAATTCCTACAGAGCAGTGAACCTGACATCTGACATCTGACATCTGACATCTGACAT 480
531 GACCAATGATGGGCTGCGAGAGATTTGCCCTGACAAAGCAAAAGCAAGACCCAGTGC 590
481 GACCAATGATGGGCTGCGAGAGATTTGCCCTGACAAAGCAAAAGCAAGACCCAGTGC 540
591 CGCTGCAGCGGGAATGTTCTGCTGCTGGGCGCTGAGGTACACACTGCGAGCTA 650
541 CGCTGCAGCGGGAATGTTCTGCTGCTGGGCGCTGAGGTACACACTGCGAGCTA 600
651 CTTTCTGACTGCGGCTGCGACTGAAGCGAGCTCAAAAGTGAAGTTGGGAAGGTTAC 710
601 CTTTCTGACTGCGGCTGCGACTGAAGCGAGCTCAAAAGTGAAGTTGGGAAGGTTAC 660
711 AACCACTGCTGCGCTGCGAGAGGAGGCAATTCACAGATCTCTGCTCCCGACGCGCG 770
661 AACCACTGCTGCGCTGCGAGAGGAGGCAATTCACAGATCTCTGCTCCCGACGCGCG 720
771 TGCAGCGCCACACAGGTGTGAGAACCAAGTCTGTGGAGGCACTCCAGGCACTGCG 830
721 TGCAGCGCCACACAGGTGTGAGAACCAAGTCTGTGGAGGCACTCCAGGCACTGCG 780
831 CAATCGACAGCAACTGCAAAAATTCATTAGAGCACTGCCCGCAGAGATGTCA----- 884

Db 781 CAGTCGACACAACTGCAGAAAATCCATTAGAGCCACTGCCCGCAGAGATGTCAAGAAC 840
OY 885 ----- 884
Db 841 ATGCGATGCTGGCGGTTCTGTGCTGCGACTGGCCTTTCTTCTGCTCTGCGACGCTTTC 900
OY 885 ----- 884
Db 901 TCTGCACTGTGAAGAGCACCCTTCTCTGTCAGAGAAACCTGGATGCTCTCAAGAG 960
OY 903 GGTCCGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 962
Db 961 GGTCCGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
OY 963 CCATCTCTCCCTGACTGCTGTGTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1022
Db 1021 CCATCTCTCCCTGACTGCTGTGTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1080
OY 1023 TCCACTGGGCTCCCGCAGGCCAGTTTGGAGGCAAGGGGTGCCGAACAGCAGAGTCT 1082
Db 1081 TCCACTGGGCTCCCGCAGGCCAGTTTGGAGGCAAGGGGTGCCGAACAGCAGAGTCT 1140
OY 1083 CTGACCTGACACAGGAGCGCGAGTTGGAACCGGAGGAGAGAGAGGAGGAGGAGGAGGAG 1142
Db 1141 CTGACCTGACACAGGAGCGCGAGTTGGAACCGGAGGAGAGAGAGGAGGAGGAGGAGGAG 1200
OY 1143 ACCAATGGCAATCTGTCACCGGCGGCTATGACTATCTGAGCAATCTGATCTAC 1202
Db 1201 ACCAATGGCAATCTGTCACCGGCGGCTATGACTATCTGAGCAATCTGATCTAC 1260
OY 1203 AATGAGCACTACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1262
Db 1261 AATGAGCACTACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
OY 1263 CCAATACCCATTCCTCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1322
Db 1321 CCAATACCCATTCCTCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
OY 1323 GAAGATGGCAAGGCTTGGCACTGAGCGAGACAGAGCAGTGTGTCACACCTCTAAC 1382
Db 1381 GAAGATGGCAAGGCTTGGCACTGAGCGAGACAGAGCAGTGTGTCACACCTCTAAC 1440
OY 1383 AGGAGCCCAAGAAATTTATCACCCATGATGATGATGATGATGATGATGATGATGATGAT 1442
Db 1441 AGGAGCCCAAGAAATTTATCACCCATGATGATGATGATGATGATGATGATGATGATGAT 1500
OY 1443 AGGAGGCGCAAGAGGACCTTCTCCCTGAGGCTGCTGCGACGTTGGATTCAAGAG 1502
Db 1501 AGGAGGCGCAAGAGGACCTTCTCCCTGAGGCTGCTGCGACGTTGGATTCAAGAG 1560
OY 1503 GCCTGATGAGGCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1562
Db 1561 GCCTGATGAGGCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
OY 1563 GCAAGTGGGCACTGGCTGGGTACGAGTGGCTTCCAGAGACTTCCCTACTGCTGAGCA 1622
Db 1621 GCAAGTGGGCACTGGCTGGGTACGAGTGGCTTCCAGAGACTTCCCTACTGCTGAGCA 1680
OY 1623 ACCTGAGGCTCTCCCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1682
Db 1681 ACCTGAGGCTCTCCCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
OY 1683 AGGAGCAGATATACCAAGTGTGCGCAGTACAGAGGCGGAGGAGGAGGAGGAGGAGGAG 1742
Db 1741 AGGAGCAGATATACCAAGTGTGCGCAGTACAGAGGCGGAGGAGGAGGAGGAGGAGGAG 1800
OY 1743 GAGGCGGCCACAGGTCACCTTGCAAGAGGTCACGAGGCGGCGCTTAAAGATTGCGTGC 1802
Db 1801 GAGGCGGCCACAGGTCACCTTGCAAGAGGTCACGAGGCGGCGCTTAAAGATTGCGTGC 1860
OY 1803 TCATCCCAAGCTTCAGAGACCTTTGGGGTTCCACACTTACGTTGAGCTAGAGTACAC 1862
Db 1861 TCATCCCAAGCTTCAGAGACCTTTGGGGTTCCACACTTACGTTGAGCTAGAGTACAC 1920

QY 351 GACGAGAAAGAAATATATAGAGCCCGACAGCCGATCTGCTGCCGTCCGCCCA 410
 Db 301 GACGAGAAAGAAATATATAGAGCCCGACAGCCGATCTGCTGCCGTCCGCCCA 360
 QY 411 GGCACCTATGTCTCAGCTAAATGTAGCCCATCCGGAGACAGTTTGTGCCACATGTGCC 470
 Db 361 GGCACCTATGTCTCAGCTAAATGTAGCCCATCCGGAGACAGTTTGTGCCACATGTGCC 420
 QY 471 GAGAAATCTTACAGAGAGCAGTGAATACCTGACCATCTGACAGTGGCCGCCCTGT 530
 Db 421 GAGAAATCTTACAGAGAGCAGTGAATACCTGACCATCTGACAGTGGCCGCCCTGT 480
 QY 531 GACCCAGATATGGGCTCGAGAGAGTTGGCCCTGCAAGCAAGCAAGCAAGCAAGTGC 590
 Db 481 GACCCAGATATGGGCTCGAGAGAGTTGGCCCTGCAAGCAAGCAAGCAAGCAAGTGC 540
 QY 591 GCGTCCAGAGCCGGAGATGTTGTGCTGCTGGGCGCTGAGTACACACTGCGAGCTA 650
 Db 541 GCGTCCAGAGCCGGAGATGTTGTGCTGCTGGGCGCTGAGTACACACTGCGAGCTA 600
 QY 651 CTTTCTGACCTGCGCCCTGGCACTGAGCCGAGCTCAAGATGAGTTGGGAGGGTAAAC 710
 Db 601 CTTTCTGACCTGCGCCCTGGCACTGAGCCGAGCTCAAGATGAGTTGGGAGGGTAAAC 660
 QY 711 AACCACTGCTGCTGCTGCAAGAGGAGGAGCTTCCAGAAATACCTCTCCCGCAGCCCGC 770
 Db 661 AACCACTGCTGCTGCTGCAAGAGGAGGAGCTTCCAGAAATACCTCTCCCGCAGCCCGC 720
 QY 771 TGCCAGCCCGCAGCAGTGTGAGAACCAAGGTCTGGTGGAGGAGCTCCAGGAGCTGCC 830
 Db 721 TGCCAGCCCGCAGCAGTGTGAGAACCAAGGTCTGGTGGAGGAGCTCCAGGAGCTGCC 780
 QY 831 CAGTCCGACAGACACCTGCAAAAATTCATTAGAGCCACTGCCCCAGAGATGTCA----- 884
 Db 781 CAGTCCGACAGACACCTGCAAAAATTCATTAGAGCCACTGCCCCAGAGATGTCAAGAAC 840
 QY 885 ----- 884
 Db 841 ATGCTGATGCTGCGCTTCTGCTGCCACTGCGCTTCTTCTGCTCTTGCACCGCTCTC 900
 QY 885 ----- 900
 Db 901 TCCTGATCTGGAAGAGAGCAACCTCTCTCTCTGAGAGAACTGGGATGCTCTCAAGAG 960
 QY 903 CGTCCGAGGAGAGAGAGCAACCTCTCTCTCTGAGAGAGCTGGAGAGCTCCGAGAGCCCAT 962
 Db 961 CGTCCGAGGAGAGAGAGCAACCTCTCTCTGAGAGAGCTGGAGAGCTCCGAGAGCCCAT 1020
 QY 963 CCATACCTTCCCTGACTGCTGTAAGAGCCACTGCTACCCATTTCTGAGATGTTTCCAGTA 1022
 Db 1021 CCATACCTTCCCTGACTGCTGTAAGAGCCACTGCTACCCATTTCTGAGATGTTTCCAGTA 1080
 QY 1023 TCCACTGGGCTCCCGCAGCCAGCTTTTGGAGGAGGAGTGGCGCAACGAGAGGCT 1082
 Db 1081 TCCACTGGGCTCCCGCAGCCAGCTTTTGGAGGAGGAGTGGCGCAACGAGAGGCT 1140
 QY 1083 CTGAGCTGACAGGAGGAGCCGCACTTGAACCCGGGAGAGCAGAGTGGGCCAGAGT 1142
 Db 1141 CTGAGCTGACAGGAGGAGCCGCACTTGAACCCGGGAGAGCAGAGTGGGCCAGAGT 1200
 QY 1143 ACCAATGAGCTATGCTACAGCGGAGTCTATGACTATCACTGGCAACATCTACATCTAC 1202
 Db 1201 ACCAATGAGCTATGCTACAGCGGAGTCTATGACTATCACTGGCAACATCTACATCTAC 1260
 QY 1203 AATGAGCAGTACTGAGGAGAGCAGCGGCTCTGAGAGACTTCCACACTACCCCGAAGCT 1262
 Db 1261 AATGAGCAGTACTGAGGAGAGCAGCGGCTCTGAGAGACTTCCACACTACCCCGAAGCT 1320
 QY 1263 CCATACCCATTTCCGAGAGAGGAGAGCCTGGGCTCTCCCGGCTCTTACACCCAGCAG 1322
 Db 1321 CCATACCCATTTCCGAGAGAGGAGAGCCTGGGCTCTCCCGGCTCTTACACCCAGCAG 1380
 QY 1323 GAAGATGGCAAGGCTTGGCAGCTAGGAGAGACAGACACTGTGTGCCACACCTCTTAC 1382

Db 1381 GAAAGTGGCAAGGCTGTGGCACTAGCGAGACAGAGCACTGTGTGCCACACCTTAAAC 1440
 QY 1383 AGGGGCCCAAGAACCAATTATACCATATACAGTACTGAGTGTGAGAAAAAGAGAGAGA 1442
 Db 1441 AGGGGCCCAAGAACCAATTATATACCATATACAGTACTGAGTGTGAGAAAAAGAGAGAGA 1500
 QY 1443 AGGGGGGACAGAGGACACTTCTCCCTTGAGGCTGCCCCAGTGGAGTTTACAGGG 1502
 Db 1501 AGGGGGGACAGAGGACACTTCTCCCTTGAGGCTGCCCCAGTGGAGTTTACAGGG 1560
 QY 1503 GCTGAGTATGAGGCGCCGGGAGAGCAGAGCCCTTAAGGGATTAAAGGTGAGAACCTGAGA 1562
 Db 1561 GCTGAGTATGAGGCGCCGGGAGAGCAGAGCCCTTAAGGGATTAAAGGTGAGAACCTGAGA 1620
 QY 1563 GCAAGTGGGCACTGCTGGGTAGCGTCCCTCCACAGAGCTCTCCCTAGCCTGAGCAA 1622
 Db 1621 GCAAGTGGGCACTGCTGGGTAGCGTCCCTCCACAGAGCTCTCCCTAGCCTGAGCAA 1680
 QY 1623 ACCTGAGGCTCTCCCGGAGAGACCCACCCCTGGGCTGTGAGCCTGAGGAC 1682
 Db 1681 ACCTGAGGCTCTCCCGGAGAGACCCACCCCTGGGCTGTGAGCCTGAGGAC 1740
 QY 1683 AGGGCAGATGATACCACTGCTGCGCCACTACAGAGCGCGACCGAGAGCAGCAGAG 1742
 Db 1741 AGGGCAGATGATACCACTGCTGCGCCACTACAGAGCGCGACCGAGAGCAGCAGAG 1800
 QY 1743 GAGAGCCGACACAGGCTACCTGCAAGAGCTCAGGAGCCCTCTTAAGATTCGTGTGC 1802
 Db 1801 GAGAGCCGACACAGGCTACCTGCAAGAGCTCAGGAGCCCTCTTAAGATTCGTGTGC 1860
 QY 1803 TCATCCCAAGCTTTCAGAGACCTTTGGGTTTCAACACTTACAGTACTGAGAGTAC 1862
 Db 1861 TCATCCCAAGCTTTCAGAGACCTTTGGGTTTCAACACTTACAGTACTGAGAGTAC 1920
 QY 1863 CTCATGAGATGAATTTATAGGAGAGAGCTCTCCCTCCCTGAGAGAGAGAGA 1922
 Db 1921 CTCATGAGATGAATTTATAGGAGAGAGCTCTCCCTCCCTGAGAGAGAGAGA 1980
 QY 1923 AGGAGTCAATTAACAAGTGGGGTGGGTAGAGATTCATAGAGAGAGAGAGTGG 1982
 Db 1981 AGGAGTCAATTAACAAGTGGGGTGGGTAGAGATTCATAGAGAGAGAGAGTGG 2040

RESULT 4
 AX32212 2136 bp DNA linear PAT 09-JAN-2002
 LOCUS AX32212
 DEFINITION Sequence 2721 from Patent W00194629.
 ACCESSION AX32212
 VERSION AX32212.1 GI:18122846
 KEYWORDS

SOURCE

human:

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
 Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
 Horrigan, S., Soppet, D.R. and Weaver, Z.
 Cancer gene determination and therapeutic screening using signature
 gene sets
 Patent: WO 0194629-A 2721 13-DEC-2001;

JOURNAL Avalon Pharmaceuticals (US)
 TITLE location/Qualifiers
 FEATURES 1..2136

source /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 446 a 706 c 608 g 376 t
 ORIGIN

Query Match 91.3%; Score 1809.2; DB 6; Length 2136;
 Best Local Similarity 94.6%; Pred. No. 0;
 Matches 1929; Conservative 0; Mismatches 3; Indels 108; Gaps 1;

QY 688 AAGATGAAGTTGGAGAGGTAACAACACACTGCTGCCCTGCAAGGACGACCTTCCAGA 747
 Db 661 AAGATGAAGTTGGAGAGGTAACAACACACTGCTGCCCTGCAAGGACGACCTTCCAGA 720
 QY 748 ATACTCTCTCCCGAGCGCCGGTCCGACCCACACAGAGTGTGAGAACCAAGTCTGG 807
 Db 721 ATACTCTCTCCCGAGCGCCGGTCCGACCCACACAGAGTGTGAGAACCAAGTCTGG 780
 QY 808 TGGAGGACGCTCCAGGACACTGCCAGTCCGACACAACTGCAAAAATCATAGAGCCAC 867
 Db 781 TGGAGGACGCTCCAGGACACTGCCAGTCCGACACAACTGCAAAAATCATAGAGCCAC 840
 QY 868 TGCCCCCAGAGATGCA----- 884
 Db 841 TGCCCCCAGAGATGCTAGAGAACCAATGCTGATGTGGCGCTTCTGTGCTGACCTGGCCTTCT 900
 QY 885 ----- 884
 Db 901 TTCTGCTCTTGGCAACCGCTTCTCTGATCTGAGAGAGCCACCTTCTCTCTGACAGA 960
 QY 885 -----GATCGCTGCTCAAGAGGCGTCCGAGGAGAGGAGCCCAATCTGAGCTGGAA 939
 Db 961 AACTGGGATCCGCTGCTCAAGAGGCGTCCGAGGAGAGGAGCCCAATCTGAGCTGGAA 1020
 QY 940 GCTGGGACCTCCGAGAGCCCATCATCTCTGACTGCTGACAGCCACTGCTACCCA 999
 Db 1021 GCTGGGACCTCCGAGAGCCCATCATCTCTGACTGCTGACAGCCACTGCTACCCA 1080
 QY 1000 TTTCTGGAGATGTTTCCCACTATTCACCTGGGCTCCCGCAGCCCAAGTTTGGAGGAG 1059
 Db 1081 TTTCTGGAGATGTTTCCCACTATTCACCTGGGCTCCCGCAGCCCAAGTTTGGAGGAG 1140
 QY 1060 GGGTTCGCCCAACAGAGAGAGTCTCTGAGCTGACAGGAGGAGCCGAGTTGGAAACCCGGGG 1119
 Db 1141 GGGTTCGCCCAACAGAGAGAGTCTCTGAGCTGACAGGAGGAGCCGAGTTGGAAACCCGGGG 1200
 QY 1120 AGCAGAGCCAGGTGGCCCGACGATACCAATGSCATTCATGTACCGCGCGGTCTATGACTA 1179
 Db 1201 AGCAGAGCCAGGTGGCCCGACGATACCAATGSCATTCATGTACCGCGCGGTCTATGACTA 1260
 QY 1180 TCACGTGGCAACATCTACATCTACATGAGCAAGTACTGGGGGAGACCCGGGTCTGTGAG 1239
 Db 1261 TCACGTGGCAACATCTACATCTACATGAGCAAGTACTGGGGGAGACCCGGGTCTGTGAG 1320
 QY 1240 ACCTCCAGCTACCCCGAACCTCCATACCCCATTCGCCGAGAGGGGGAGACCTGGCCCTC 1299
 Db 1321 ACCTCCAGCTACCCCGAACCTCCATACCCCATTCGCCGAGAGGGGGAGACCTGGCCCTC 1380
 QY 1300 CCGGGCTCTTACACCCCGACAGAGAGTGGCAAGGCTTGGCACCCTAGCGAGACAGAC 1359
 Db 1381 CCGGGCTCTTACACCCCGACAGAGAGTGGCAAGGCTTGGCACCCTAGCGAGACAGAC 1440
 QY 1360 ACTGTGTGACACACCTCTTAAACAGGGGCCCAAGAGAACCAATTATACCCATGACTGAC 1419
 Db 1441 ACTGTGTGACACACCTCTTAAACAGGGGCCCAAGAGAACCAATTATACCCATGACTGAC 1500
 QY 1420 TGAGTCTGAGAAAAGGAGGAAGAGGGGGGACAAAGGGGACCTTCTGAGAGGCTGCC 1479
 Db 1501 TGAGTCTGAGAAAAGGAGGAAGAGGGGGGACAAAGGGGACCTTCTGAGAGGCTGCC 1560
 QY 1480 CTGCCACAGTGGGATTCACAGGGGGCTGAGTAGGGCCCGGGGAAAGCAAGGCTTAAGGGA 1539
 Db 1561 CTGCCACAGTGGGATTCACAGGGGGCTGAGTAGGGCCCGGGGAAAGCAAGGCTTAAGGGA 1620
 QY 1540 TTAAGGCTCAACACCTCTGAGAGAGTGGGACCTGGCTGGGTAGGTTGCCCTTCACAG 1599
 Db 1621 TTAAGGCTCAACACCTCTGAGAGAGTGGGACCTGGCTGGGTAGGTTGCCCTTCACAG 1680
 QY 1600 GACTTCTCTACTGCTAGCAAAACCTGAGGCTCCCGGAGAGCCACACCCCTCTGGG 1659
 Db 1681 GACTTCTCTACTGCTAGCAAAACCTGAGGCTCCCGGAGAGCCACACCCCTCTGGG 1740
 QY 1660 GCTGCTCAGCTCAGGCAAGGAGCAATGATATCCAACTGCTGCCACTAGACAGC 1719

Db 1741 GCTGCTCAGCTCAGGACAGGACAGGAGCAGATGATACCAATGCTGCCACTAGCGACG 1800
 QY 1720 CCGCAGCGGAGACAGGACAGGAGAGCGCCACACGGGTACACGAGCAAGAGAGTCAAGG 1779
 Db 1801 CCGCAGCGGAGACAGGACAGGAGAGCGCCACACGGGTACACGAGCAAGAGAGTCAAGG 1860
 QY 1780 CCCCTCTAAAGATTCGTGTGCTCATCCCAAGCTTCAGAGACCCTTGGGGTTCCACA 1839
 Db 1861 CCCCTCTAAAGATTCGTGTGCTCATCCCAAGCTTCAGAGACCCTTGGGGTTCCACA 1920
 QY 1840 CTTACAGTGGACTGAGTAGACCTCGCATGAGATGAATATATAGGAGAGAGCTCTTC 1899
 Db 1921 CTTACAGTGGACTGAGTAGACCTCGCATGAGATGAATATATAGGAGAGAGCTCTTC 1980
 QY 1900 CTTCCCTCCCTAGAGGAGAGGAGGAGGATTAACAACTAGAGGGGTGGGTAGAGATTC 1959
 Db 1981 CTTCCCTCCCTAGAGGAGAGGAGGAGGATTAACAACTAGAGGGGTGGGTAGAGATTC 2040
 QY 1960 CTAGGTATGGGAGAGATTTGG 1982
 Db 2041 CTAGGTATGGGAGAGATTTGG 2063

RESULT 3
 AX331906 2136 bp DNA linear PAT 09-JAN-2002
 LOCUS
 DEFINITION Sequence 2415 from Patent W00194629.
 ACCESSION AX331906
 VERSION AX331906.1 GI:18122540
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1
 AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
 Horrigan,S., Soppet,D.R. and Weaver,Z.
 TITLE Cancer gene determination and therapeutic screening using signature
 gene sets
 JOURNAL Patent: WO 0194629-A 2415 13-DEC-2001;
 Avalon Pharmaceuticals (US)
 FEATURES
 source 1..2136
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 446 a 706 c 608 g 376 t
 ORIGIN

Query Match 91.3% Score 1809.2; DB 6; Length 2136;
 Best Local Similarity 94.6% Pred. No. 0;
 Matches 1929; Conservative 0; Mismatches 3; Indels 108; Gaps 1;
 QY 51 GCCCTGAGAGCCCGGCTGCGGCTCCCGGCTGAGGCTGACATCGGCTCGAGTCCG 110
 Db 1 GCCCTGAGAGCCCGGCTGCGGCTCCCGGCTGAGGCTGACATCGGCTCGAGTCCG 60
 QY 111 TCCAGAGCTGTGGGCTCGGGACAGCCGCGCACGCTGCCAGAGAGTGGGCTCTCTGC 170
 Db 61 TCCAGAGCTGTGGGCTCGGGACAGCCGCGCACGCTGCCAGAGAGTGGGCTCTCTGC 120
 QY 171 CTTCTCCAGAGCCCGGCTGCTGAGGCGGCTGGGAGTGGGCGGACATCTCTGCTG 230
 Db 121 CTTCTCCAGAGCCCGGCTGCTGAGGCGGCTGGGAGTGGGCGGACATCTCTGCTG 180
 QY 231 TGGGCACTCTGCCCCCGGCTGAGGGGCTCTGTGTGTGGGCTCTTGGGCTC 290
 Db 181 TGGGCACTCTGCCCCCGGCTGAGGGGCTCTGTGTGTGGGCTCTTGGGCTC 240
 QY 291 CTGGCAGATGGAGCCCGGAGGCTGCTCCATATGTGTGGAGAACCAAGCTGTGAG 350
 Db 241 CTGGCAGATGGAGCCCGGAGGCTGCTCCATATGTGTGGAGAACCAAGCTGTGAG 300

```

Db 1382 GCGCCCGGGGGCTCTTAACACCCACAGGAGATGGCAAGCTTGGCAGCTAGCGGAGA 1441
QY 1354 CAGAGACTGTGTGGCCACACCTCTTAACAGAGGGGCCAAGAACCAATTATTCACCATG 1413
Db 1442 CAGAGACTGTGTGGCCACACCTCTTAACAGAGGGGCCAAGAACCAATTATTCACCATG 1501
QY 1414 ACTGACTGTGTGTGAAAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1473
Db 1502 ACTGACTGTGTGTGAAAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1551
QY 1474 GCTGCGCGCGCGAGTGGGATTCACAGGGGCTGATAGGCGCCGGGAGACAGAGCCCT 1533
Db 1562 GCTGCGCGCGCGAGTGGGATTCACAGGGGCTGATAGGCGCCGGGAGAGAGAGCCCT 1621
QY 1534 AAGGATTAAGGCTCAGACACCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1593
Db 1522 AAGGATTAAGGCTCAGACACCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1681
QY 1594 CCACAGACTCTCCCTACTGCTGAGCAAACTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1653
Db 1682 CCACAGACTCTCCCTACTGCTGAGCAAACTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1741
QY 1654 CCTGGGGGTGCTGACCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1713
Db 1742 CCTGGGGGTGCTGACCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1801
QY 1714 AGCAGCGGAGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1773
Db 1802 GGCAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1861
QY 1774 CAGGAGCGGCTCTTAAGAGATTCGTGTGTCATCCCAAGCTTCAGAGAGAGAGAGAG 1833
Db 1862 CAGGAGCGGCTCTTAAGAGATTCGTGTGTCATCCCAAGCTTCAGAGAGAGAGAGAG 1921
QY 1834 TCACACTTACGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1893
Db 1922 TCACACTTACGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1981
QY 1894 TCCTTCCTCCCTCTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1953
Db 1982 TCCTTCCTCCCTCTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2041
QY 1954 GGATTCCTAGATGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1982
Db 2042 GGATTCCTAGATGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2070

RESULT 2
AK027080 2148 bp mRNA linear PRI 29-SEP-2000
LOCUS AK027080
DEFINITION Homo sapiens cDNA: FLJ23427 fls, clone HRC04788, highly similar to
HUMNFRP Homo sapiens tumor necrosis factor receptor 2 related
protein mRNA.
ACCESSION AK027080
VERSION AK027080.1 GI:1044011
KEYWORDS oligo capping; fls (full insert sequence).
SOURCE Homo sapiens primary human renal epithelial cells cDNA to mRNA,
clone.lib:HRC clone:HRC04788.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (sites)
AUTHORS Kawabata,A., Hiki,J., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
NEO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2148)
AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,

```

```

COMMENT
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
NEO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan: cDNA full insert
sequencing: Research Association for Biotechnology: cDNA library
construction, 5'- & 3'-end one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).

FEATURES
source
1. 2148
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HRC04788"
/cell_type="primary human renal epithelial cells"
/clone_lib="HRC"
/note="cloning vector pME18SFL3"
misc_feature
1. 2148
/note="highly similar to HUMNFRP Homo sapiens tumor
necrosis factor receptor 2 related protein mRNA"
BASE COUNT 462 a 705 c 606 g 375 t
ORIGIN
Query Match 92.5%; Score 1833.8; DB 9; Length 2148;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 1953; Conservative 0; Mismatches 2; Indels 108; Gaps 1;

QY 28 ACTTCTGAGCTCCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 87
Db 1 ACTTCTGAGCTCCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
QY 88 GTGCAATATGGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 147
Db 61 GTGCAATATGGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 148 GCCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 207
Db 121 GCCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 208 GAGTGGCGCCATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 267
Db 181 GAGTGGCGCCATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 268 TGGTGTGGGCTCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 327
Db 241 TGGTGTGGGCTCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 328 CGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 387
Db 301 CGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 388 TCTGCTGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 447
Db 361 TCTGCTGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 448 AACAGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 507
Db 421 AACAGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 508 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 567
Db 481 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 568 CAAGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 627
Db 541 CAAGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 628 TCGATGTACACATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 687
Db 601 TCGATGTACACATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660

```

JOURNAL

Submitted (02-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Email: cgabs-remail.nih.gov
Tissue Procurement: CLONTECH

CNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLI)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mdickpaxil.stanford.edu
Dickson, M., Schmitz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNLI at: <http://image.llnl.gov>.
Series: ILNLI Plate: 37 Row: b Column: 13
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA g1: 4505038.
Location/Qualifiers

FEATURES

1. 2161

/organism="Homo sapiens"
/db_xref="LocusID:4055"
/db_xref="taxon:9606"
/clone="MGC:22593 IMAGE:4703362"
/tissue_type="Lung"
/clone_lib="NIH_MGC_77"
/lab_host="DH10B"
/note="Vector: PDNR-LIB"
199. 1506
/codon_start=1
/product="Lymphotoxin beta receptor (TNFR superfamily, member 3)."
/protein_id="AAH26262.1"
/db_xref="GI:20072213"
/translation="MLPWATSPAGLAWGPELVGLFGLAASQAPVAPYASENQTCR
DDEKEYEDPHRICKSRCPGTIVSAKCSIRIDPTVCATAEINSHMNYLTQCLR
PCDPVAGLEIAIPCTSKRKTCRCQPMFCAMALCTCELSIDCPGTEALIKDEY
GKNNHCVCRAKHPQNTSSPARCOPHTRCENQGIYEAAPGASDPTCKNLEPLP
PEMSCTMLAVILPLAFLLIATVFSCTIKSPISCRKIGSLILKRPGEENPVAG
SMPPKAAHYFPDLPVPLPISGDVSPVSTGLPAALVAGVQDQSPDLREPDLE
PGEQSVAAHTNGIHTVGSMTITGNIYINGVLTGAGPGGDLPATPPPPYPIPEEG
DGGPPELSTPHQEDGKAMHLETHCGAPSPNRPNGRPTTHD"

CDS

BASE COUNT 473 a 708 c 606 g 374 t
ORIGIN

Query Match 92.8%; Score 1839.8; DB 9; Length 2161;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 1959; Conservative 0; Mismatches 2; Indels 108; Gaps 1;

22 AGCCCACTCTGAGCTCGCCATGGAGGAGCCCTGGAGGCGCCGCTGCGCTCCGGC 81
Db 2 ACTCCCACTCTGAGCTCGCCATGGAGGAGCCCTGGAGGCGCCGCTGCGCTCCGGC 61
82 CTTGGGGTGACATCGGCTCTGAGTCCCTCCAGAGGCTTGGGCTGGGAGCGCGCC 141
Db 62 CTTGGGGTGACATCGGCTCTGAGTCCCTCCAGAGGCTTGGGCTGGGAGCGCGCC 121
142 ACCGCTGCGCAGAGAGCTGGGCTCTGCTTCTCCAGGCGCCCAAGTTGTGGCCG 201
Db 122 ACCGCTGCGCAGAGAGCTGGGCTCTGCTTCTCCAGGCGCCCAAGTTGTGGCCG 181
202 CTGGCGAGTGGCGGCTCTGCTTCTGGGCGCACCTTGGCGCGCGCTGCTGGG 261
Db 182 CTGGCGAGTGGCGGCTCTGCTTCTGGGCGCACCTTGGCGCGCGCTGCTGGG 241
262 GGCCTCTGTGTGGGCTCTTGGGCTCTGTGACAGATGAGACCCCAAGGCTGCTC 321
Db 242 GGCCTCTGTGTGGGCTCTTGGGCTCTGTGACAGATGAGACCCCAAGGCTGCTC 301

322 CATATGGCTCGAGAGACCACTGACAGGAGCCAGAGAAAGAAATGATGAGCCCGC 381
Db 302 CATATGGCTCGAGAGACCACTGACAGGAGCCAGAGAAAGAAATGATGAGCCCGC 361
382 ACCGATCTGCTGCTCCCGCTGCGCCGAGGACCTATGTCTACGTAATGATGCGCA 441
Db 362 ACCGATCTGCTGCTCCCGCTGCGCCGAGGACCTATGTCTACGTAATGATGCGCA 421
442 TCGGGGACACGTTGTGGCCACATGTCGCGGAAATTCCTACAGACGACCTGGAAC 501
Db 422 TCGGGGACACGTTGTGGCCACATGTCGCGGAAATTCCTACAGACGACCTGGAAC 481
502 TGACATCTGACAGCTGTCGCGCCCTGTGACCCAGATGATGGGCTGAGAGATTCGC 561
Db 482 TGACATCTGACAGCTGTCGCGCCCTGTGACCCAGATGATGGGCTGAGAGATTCGC 541
562 CTGACACAGCAAGCAAGAGACCAAGTGCCTGTCGACCGGGAATTTGTGCTGCT 621
Db 542 CTGACACAGCAAGCAAGAGACCAAGTGCCTGTCGACCGGGAATTTGTGCTGCT 601
622 GGGCTCTGAGTGTACACACTGGGAGCTTCTGTGACTGCGCGCTGGACCTGAAACCG 681
Db 602 GGGCTCTGAGTGTACACACTGGGAGCTTCTGTGACTGCGCGCTGGACCTGAAACCG 661
682 AGCTCAAGATGAAGTTGGGAAGGGTAACAACACACTGCTCCCTGCAAGGACAGGCACT 741
Db 662 AGCTCAAGATGAAGTTGGGAAGGGTAACAACACACTGCTCCCTGCAAGGACAGGCACT 721
742 TCCAGATTAATCTCTCTCCCGCCAGGCGCGCTGACAGGCGCCACACAGGTGTGAGAACCAAG 801
Db 722 TCCAGATTAATCTCTCTCCCGCCAGGCGCGCTGACAGGCGCCACACAGGTGTGAGAACCAAG 781
802 GTCTGTGGAGAGGAGCTCCAGAGGCTCCAGTCCGACACAACTGGAAAAATCATTTAG 861
Db 782 GTCTGTGGAGAGGAGCTCCAGAGGCTCCAGTCCGACACAACTGGAAAAATCATTTAG 841
862 AGCCACTGCCCCAGAGATGTCA----- 884
Db 842 AGCCACTGCCCCAGAGATGTCA----- 864
885 ----- 884
902 CTTTCTTTCCTCTTCCACCGCTTCTCTGACATCTGGAAGAGCCACCTTCTCTCT 961
Db 885 ----- 901
962 GCAAGAACTGGGATCTCTCTCTCAGAGGCTCCGAGGAGAGGAGCCCAATCTCTAG 1021
934 CTGAAAGCTGGAGGCTCCGAAAGGCGCCATCACTTCTCTGACTTGGTACAGCACTGC 993
Db 1022 CTGAAAGCTGGAGGCTCCGAAAGGCGCCATCACTTCTCTGACTTGGTACAGCACTGC 1001
994 TACCAATTTTGGAGATGTTTCCCACTATCCACTGAGGCTCCCGACAGCCCAAGTTTGG 1053
Db 1082 TACCAATTTTGGAGATGTTTCCCACTATCCACTGAGGCTCCCGACAGCCCAAGTTTGG 1141
1054 AGGAGAGGCTGGCGCAACACAGAGTCTCTGACCTGAGACCTGAGAGAGCCGACATGGAAC 1113
Db 1142 AGGAGAGGCTGGCGCAACACAGAGTCTCTGACCTGAGACCTGAGAGAGCCGACATGGAAC 1201
1114 CCGGGGAGACAGACCAAGTGGTCCAGGCTACCAATGATGATGACCGGGGGTCTA 1173
Db 1202 CCGGGGAGACAGACCAAGTGGTCCAGGCTACCAATGATGATGACCGGGGGTCTA 1261
1262 TGACTATCACTGCAACATCTACATCTACAAATGAGACAGTATGAGGGGGACACCGGGTTC 1321
1234 CTGGAGACCTCCAGCTACACCGGCAACCTGCAATCCCATTTCCGAAAGAGGGGACCTTC 1293
Db 1322 CTGGAGACCTCCAGCTACACCGGCAACCTGCAATCCCATTTCCGAAAGAGGGGACCTTC 1381
1294 GGCCTCCCGGCTCTCTACACCCACAGAGAGATGGCAAGGCTTGGCACTAGCGGAGA 1353

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:04:15 ; Search time 3695.99 Seconds
(without alignments)
15606.600 Million cell updates/sec

Title: US-09-917-372-2

Perfect score: 1982

Sequence: 1 gccccgcgcgcagctcgtc.....ggtatggggaagagtttgg 1982

Scoring table:

IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_hlg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_to: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vl: *
30: em_hlg_hum: *
31: em_hlg_inv: *
32: em_hlg_other: *
33: em_hlg_mus: *
34: em_hlg_pln: *
35: em_hlg_rtd: *
36: em_hlg_mam: *
37: em_hlg_vrt: *
38: em_sy: *
39: em_hlg_hum: *
40: em_hlg_mus: *
41: em_hlg_other: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1839.8	92.8	2161	9	BC026262
2	1833.8	92.5	2148	9	AK027080
3	1809.2	91.3	2136	6	AX331906
4	1809.2	91.3	2136	6	AX332212
5	1809.2	91.3	2136	6	AX409488
6	1809.2	91.3	2136	6	HMTNFRRP
7	1732.2	87.4	2091	9	AK095418
8	841.4	42.5	140026	9	AC005840
9	765.8	38.6	2076	10	MMU25173
10	673.8	34.0	1614	10	MUSLYMPHOB
11	444.6	22.4	527	6	AX381072
12	296.8	15.0	400	11	G11431
13	281.8	14.2	178228	2	AC128082
14	264	13.3	187998	2	AC125909
15	198	10.0	1605	9	HMTNFRNEC
16	178.8	9.0	250	11	G11535
17	151.2	7.6	187998	2	AC125909
18	149.6	7.5	2048	10	MMU30798
19	99	5.0	323	9	HS2789R
20	97.6	4.9	193	9	HS2789R
21	96.8	4.9	178228	2	AC128082
22	73.4	3.7	125020	9	AF429315
23	72	3.6	1557	6	AX054825
24	72	3.6	1557	6	I36197
25	72	3.6	1641	6	I36196
26	72	3.6	2224	6	ARI52033
27	72	3.6	2253	6	A78517
28	72	3.6	2394	9	HMTNFRIT
29	72	3.6	3380	11	G26865
30	72	3.6	3683	6	AX333705
31	72	3.6	3683	6	AX348016
32	72	3.6	3683	6	AX348018
33	72	3.6	3683	6	AX348020
34	72	3.6	3683	9	HMTNFR
35	69.4	3.5	125020	9	AF429315
36	62.6	3.2	691	6	I36350
37	62.6	3.2	705	6	ARI74443
38	62.4	3.1	579	6	ARI68180
39	62.4	3.1	591	6	ARI68183
40	62.4	3.1	852	9	AF373877
41	62.4	3.1	852	9	AF373878
42	62.4	3.1	1596	6	ARI68182
43	62.4	3.1	1621	9	BC002794
44	62.4	3.1	1724	6	ARI69913
45	62.4	3.1	1724	6	ARI72617

ALIGNMENTS

RESULT 1
LOCUS BC026262 2161 bp mRNA linear PRI 08-APR-2002
DEFINITION Homo sapiens, lymphotoxin beta receptor (TNFR superfamily, member 3), clone MGC:22593 IMAGE:4703362, mRNA, complete cds.
ACCESSION BC026262
VERSION BC026262.1 GI:20072212
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2161)
AUTHORS Strausberg, R.
TITLE Direct Submission

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:04:15 ; Search time 792.53 Seconds

(without alignments)
15606.600 Million cell updates/sec

Title: US-09-917-372-2_COPY_853_1277

Perfect score: 425
Sequence: 1 atccattagagccactgcc.....aacctccattaccattccc 425Scoring table:
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_da:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pln:*
35: em_hlg_rtd:*
36: em_hlg_mam:*
37: em_hlg_vrt:*
38: em_sy:*
39: em_hlgo_hum:*
40: em_hlgo_mus:*
41: em_hlgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	393	92.5	2091	9 AK095418	AK095418 Homo sapi
2	393	92.5	2136	6 AX331906	AX331906 Sequence
3	393	92.5	2136	6 AX332212	AX332212 Sequence
4	393	92.5	2136	6 AX409488	AX409488 Sequence
5	393	92.5	2136	6 HUMT98RR	HUMT98RR Sequence
6	393	92.5	2148	9 AK027080	AK027080 Homo sapi
7	393	92.5	2161	9 BC026262	BC026262 Homo sapi
8	235.4	55.4	140026	9 AC005840	AC005840 Homo sapi
9	219.2	51.6	1614	10 MUSLYMPHOB	MUSLYMPHOB
10	219.2	51.6	2076	10 MMU29173	MMU29173 Mus musculu
11	117.6	27.7	178228	2 AC128082	AC128082 Rattus no
12	82.4	19.4	187998	2 AC125909	AC125909 Rattus no
13	40	9.4	125020	9 AF429315	AF429315 Homo sapi
14	38.4	9.0	123020	9 AF429315	AF429315 Homo sapi
15	37.8	8.9	6349	5 GGTNY	GGTNY
16	37.8	8.9	61851	9 AC022240	AC022240 Homo sapi
17	37.8	8.9	159950	2 AC016413	AC016413 Homo sapi
18	37.8	8.9	179357	9 AC074191	AC074191 Homo sapi
19	37.4	8.8	66808	1 SAU421825	SAU421825 Stigmatel
20	36.8	8.7	232957	2 AC106606	AC106606 Rattus no
21	36.6	8.6	151437	2 AC118809	AC118809 Rattus no
22	36	8.5	170302	2 HST402N21	HST402N21 Human DNA
23	36	8.5	187336	2 AP005427	AP005427 Oryza sat
24	35.6	8.4	62496	2 AC090664	AC090664 Homo sapi
25	35.6	8.4	97120	9 AP004147	AP004147 Homo sapi
26	35.6	8.4	102173	9 AP000796	AP000796 Homo sapi
27	35.6	8.4	221357	2 AP004246	AP004246 Homo sapi
28	35.4	8.3	60660	9 AC005903	AC005903 Homo sapi
29	35.4	8.3	64784	2 AC036134	AC036134 Homo sapi
30	35.4	8.3	135622	9 AC093128	AC093128 Papio cyn
31	35.4	8.3	207541	2 AC121252	AC121252 Homo sapi
32	35.4	8.3	218687	2 AC079501	AC079501 Mus muscu
33	35.4	8.3	218753	2 AC079515	AC079515 Mus muscu
34	35.2	8.3	333	6 AX002166	AX002166 Sequence
35	35.2	8.3	333	9 HSY12721	HSY12721 H.sapiens m
36	35.2	8.3	36631	9 HUMCOL7A1X	L23982 Homo sapien
37	35.2	8.3	164407	2 AC097229	AC097229 Sus scrofa
38	35.2	8.3	232784	2 AC113522	AC113522 Mus muscu
39	35	8.2	1336	3 AF350267	AF350267 Argiope t
40	35	8.2	127178	9 AC005160	AC005160 Homo sapi
41	35	8.2	146170	2 AC079224	AC079224 Homo sapi
42	35	8.2	153570	2 AC027307	AC027307 Homo sapi
43	35	8.2	230372	2 AC073693	AC073693 Mus muscu
44	34.8	8.2	936	9 AF143881	AF143881 Homo sapi
45	34.8	8.2	6150	10 MMU76618	MMU76618 Mus musculu

ALIGNMENTS

RESULT 1
LOCUS AK095418 2091 bp mRNA linear PRI 15-JUL-2002
DEFINITION Homo sapiens cDNA FLJ38099 fis, clone D3OST1000238, highly similar
to LYMPHOTOXIN-BETA RECEPTOR PRECURSOR.
ACCESSION AK095418
VERSION AK095418.1 GI:21754669
KEYWORDS Oligo capping; fis (full insert sequence).
SOURCE Homo sapiens CD34+ Cells cDNA to mRNA, clone_1b:D3OST1
clone:D3OST1000238.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Oshima,A., Takahashi-Fujii,A., Tanase,T., Imose,N., Takeuchi,K.,

Arta, M., Musashino, K., Yunki, H., Hara, H., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Makatsui, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Matsumura, M., Murakawa, K., Kanehori, K., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuko, Y., Nagai, K. and Isogai, T.

NEO human cDNA sequencing project

Unpublished
2 (bases 1 to 2091)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB) (supported by Japan Construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'-6,3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES
source
1. 2091
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="D3OST1000238"
/cell_type="CD34+ Cells"
/clone_id="D3OST1"
/note="cloning vector: PME18SFL3-mRNA from CD34+ cells after 3-days ODF induction. Primary culture, CD34+ cells"

BASE COUNT
440 a 691 c 594 g 366 t

ORIGIN
Query Match 92.5%; Score 393; DB 9; Length 2091;
Best Local Similarity 100.0%; Pred. No. 2.1e-92;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

33 GGATGCTGCTCAAGAGGCGTCCGAGAGAGAGACCAATCTGTAGCTGGAGGCTG 92
Db 930 GGATGCTGCTCAAGAGGCGTCCGAGAGAGAGACCAATCTGTAGCTGGAGGCTG 989
QY 93 GAGCCTCCGAAGGCCCATCTACTTCCCTGACTTGGTACAGCCACTGCTACCATTTCT 152
Db 990 GAGCCTCCGAAGGCCCATCTACTTCCCTGACTTGGTACAGCCACTGCTACCATTTCT 1049
QY 153 GGAGATGTTTCCCAATATCATGCTGAGTCCCGGAGCCGAGCTTTTGGAGGCAAGGCTG 212
Db 1050 GGAGATGTTTCCCAATATCATGCTGAGTCCCGGAGCCGAGCTTTTGGAGGCAAGGCTG 1109
QY 213 CCGCAACAGCAGAGTCTGACCTGACCTGACAGGAGCCGAGTGGAAACCCGGGGAGAGC 272
Db 1110 CCGCAACAGCAGAGTCTGACCTGACCTGACAGGAGCCGAGTGGAAACCCGGGGAGAGC 1169
QY 273 AGCCAGGTGGCCACGATGACCAATGATGTCACCGGGGCTCTATGACTATCACT 332
Db 1170 AGCCAGGTGGCCACGATGACCAATGATGTCACCGGGGCTCTATGACTATCACT 1229
QY 333 GGCACATCTACATCTACATGAGCAGCAAGTACGAGGAGGAGCCAGCGGGTCTCTGAGACCTC 392
Db 1230 GGCACATCTACATCTACATGAGCAGCAAGTACGAGGAGGAGCCAGCGGGTCTCTGAGACCTC 1289
QY 393 CCAGCTACCCCGAACCCTCATACCCCATTTCCC 425
Db 1290 CCAGCTACCCCGAACCCTCATACCCCATTTCCC 1322

RESULT 2
AX331906 2136 bp DNA linear PAT 09-JAN-2002
LOCUS AX331906
DEFINITION Sequence 2415 from Patent WO0194629.
ACCESSION AX331906

VERSION
AX331906.1 GI:18122540

KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
1
Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horrikan, S., Soppet, D.R. and Weaver, Z.
Cancer gene determination and therapeutic screening using signature gene sets
Patent: WO 0194629-A 2415 13-DEC-2001;
Avalon Pharmaceuticals (US)
location/Qualifiers
1. 2136
/organism="Homo sapiens"
/db_xref="taxon:9606"

FEATURES
source
BASE COUNT
446 a 706 c 608 g 376 t

ORIGIN
Query Match 92.5%; Score 393; DB 6; Length 2136;
Best Local Similarity 100.0%; Pred. No. 2.1e-92;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

33 GGATGCTGCTCAAGAGGCGTCCGAGAGAGAGACCAATCTGTAGCTGAGACTCG 92
Db 943 GGATGCTGCTCAAGAGGCGTCCGAGAGAGAGACCAATCTGTAGCTGAGACTCG 1002
QY 93 GAGCCTCCGAAGGCCCATCTACTTCCCTGACTTGGTACAGCCAGTCTACCATTTCT 152
Db 1003 GAGCCTCCGAAGGCCCATCTACTTCCCTGACTTGGTACAGCCAGTCTACCATTTCT 1062
QY 153 GGAGATGTTTCCCAATATCATGCTGAGTCCCGGAGCCGAGCTTTTGGAGGCAAGGCTG 212
Db 1063 GGAGATGTTTCCCAATATCATGCTGAGTCCCGGAGCCGAGCTTTTGGAGGCAAGGCTG 1122
QY 213 CCGCAACAGCAGAGTCTGCTGACCTGACAGGAGGAGCCGAGTGGAAACCCGGGGAGAGC 272
Db 1123 CCGCAACAGCAGAGTCTGCTGACCTGACAGGAGGAGCCGAGTGGAAACCCGGGGAGAGC 1182
QY 273 AGCCAGGTGGCCACGATGACCAATGATGTCACCGGGGCTCTATGACTATCACT 332
Db 1183 AGCCAGGTGGCCACGATGACCAATGATGTCACCGGGGCTCTATGACTATCACT 1242
QY 333 GGCACATCTACATCTACATGAGCAGCAAGTACGAGGAGGAGCCAGCGGGTCTCTGAGACCTC 392
Db 1243 GGCACATCTACATCTACATGAGCAGCAAGTACGAGGAGGAGCCAGCGGGTCTCTGAGACCTC 1302
QY 393 CCAGCTACCCCGAACCCTCATACCCCATTTCCC 425
Db 1303 CCAGCTACCCCGAACCCTCATACCCCATTTCCC 1335

RESULT 3
AX332212 2136 bp DNA linear PAT 09-JAN-2002
LOCUS AX332212
DEFINITION Sequence 2721 from Patent WO0194629.
ACCESSION AX332212
VERSION AX332212.1 GI:18122846
KEYWORDS
SOURCE
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
1
Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horrikan, S., Soppet, D.R. and Weaver, Z.
Cancer gene determination and therapeutic screening using signature gene sets
Patent: WO 0194629-A 2721 13-DEC-2001;
Avalon Pharmaceuticals (US)
location/Qualifiers
1. 2136

FEATURES
source

/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 446 a 706 c 608 g 376 t
ORIGIN

Query Match 92.5%; Score 393; DB 6; Length 2136;
Best Local Similarity 100.0%; Pred. No. 2.1e-92;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 GGATCCCTGCTCAAGAGGGGTCGGCAGAGGAGAGGCCCAATCTTGTAGTGGAGCTGG 92
DB 943 GGATCCCTGCTCAAGAGGGGTCGGCAGAGGAGAGGCCCAATCTTGTAGTGGAGCTGG 1002
QY 93 GAGCTCCGAGAGGCCCATCATCTTCCCTGACTTGTAGAGGCATGCTACCATTTTCT 152
DB 1003 GAGCTCCGAGAGGCCCATCATCTTCCCTGACTTGTAGAGGCATGCTACCATTTTCT 1062
QY 153 GGAGATGTTTCCCAATATCCACTGGGCTCCCGCAGAGGCCCAATTTTGGAGGAGGGGTG 212
DB 1063 GGAGATGTTTCCCAATATCCACTGGGCTCCCGCAGAGGCCCAATTTTGGAGGAGGGGTG 1122
QY 213 CCGCAACAGCAGAGTCTCTGACCTGACAGGAGGCCGACAGTTGGAAACCCGGGAGCAG 272
DB 1123 CCGCAACAGCAGAGTCTCTGACCTGACAGGAGGCCGACAGTTGGAAACCCGGGAGCAG 1182
QY 273 AGCCAGGTGGCCACGATGACATGTCACCGGGGGGTATGACTATCACT 332
DB 1183 AGCCAGGTGGCCACGATGACATGTCACCGGGGGGTATGACTATCACT 1242
QY 333 GGCACATCTACATCTACATGAGCAGAGTACTGGGGGAGCACCGGGTCTCTGGAGACCTC 392
DB 1243 GGCACATCTACATCTACATGAGCAGAGTACTGGGGGAGCACCGGGTCTCTGGAGACCTC 1302
QY 393 CCAGCTACCCCGCAACCTCCATACCCCATTTCC 425
DB 1303 CCAGCTACCCCGCAACCTCCATACCCCATTTCC 1335

RESULT 4
AX409488
LOCUS AX409488 2136 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 2135 from Patent WO0229103.
ACCESSION AX409488
VERSION AX409488.1 GI:21442193
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
1
Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2135 11-APR-2002;
GENE LOGIC INC (US)
FEATURES
source 1. 2136
Location/Qualifiers
1. 2136
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="EMBL/Genbank Accession No. L04270"

BASE COUNT 446 a 706 c 608 g 376 t

Query Match 92.5%; Score 393; DB 6; Length 2136;
Best Local Similarity 100.0%; Pred. No. 2.1e-92;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 GGATCCCTGCTCAAGAGGGGTCGGCAGAGGAGAGGCCCAATCTTGTAGTGGAGCTGG 92
DB 943 GGATCCCTGCTCAAGAGGGGTCGGCAGAGGAGAGGCCCAATCTTGTAGTGGAGCTGG 1002
QY 93 GAGCTCCGAGAGGCCCATCATCTTCCCTGACTTGTAGAGGCATGCTACCATTTTCT 152
DB 1003 GAGCTCCGAGAGGCCCATCATCTTCCCTGACTTGTAGAGGCATGCTACCATTTTCT 1062

QY 153 GGAGATGTTTCCCAATATCCACTGGGCTCCCGCAGAGGCCCAATTTTGGAGGAGGGGTG 212
DB 1063 GGAGATGTTTCCCAATATCCACTGGGCTCCCGCAGAGGCCCAATTTTGGAGGAGGGGTG 1122
QY 213 CCGCAACAGCAGAGTCTCTGACCTGACAGGAGGCCGACAGTTGGAAACCCGGGAGCAG 272
DB 1123 CCGCAACAGCAGAGTCTCTGACCTGACAGGAGGCCGACAGTTGGAAACCCGGGAGCAG 1182
QY 273 AGCCAGGTGGCCACGATGACATGTCACCGGGGGGTATGACTATCACT 332
DB 1183 AGCCAGGTGGCCACGATGACATGTCACCGGGGGGTATGACTATCACT 1242
QY 333 GGCACATCTACATCTACATGAGCAGAGTACTGGGGGAGCACCGGGTCTCTGGAGACCTC 392
DB 1243 GGCACATCTACATCTACATGAGCAGAGTACTGGGGGAGCACCGGGTCTCTGGAGACCTC 1302
QY 393 CCAGCTACCCCGCAACCTCCATACCCCATTTCC 425
DB 1303 CCAGCTACCCCGCAACCTCCATACCCCATTTCC 1335

RESULT 5
HUMTNFRP 2136 bp mRNA linear PRI 03-AUG-1993
LOCUS HUMTNFRP
DEFINITION Homo sapiens (clone CD18) tumor necrosis factor receptor 2 related protein mRNA, complete cds.
ACCESSION L04270.1 GI:339761
VERSION L04270.1
KEYWORDS tumor necrosis factor receptor 2 related protein.
SOURCE Homo sapiens (library: liver cDNA of P.M.) liver cDNA to mRNA.
ORGANISM Homo sapiens

REFERENCE
1
Baens, M., Chaffanet, M., Cassiman, J. J., van den Berghe, H. and Maynen, P.
TITLE Construction and evaluation of a hncDNA library of human 12p transcribed sequences derived from a somatic cell hybrid
JOURNAL Genomics 16 (1), 214-218 (1993)
MEDLINE 93252381
PUBMED 8486360
FEATURES
source 1. 2136
Location/Qualifiers
1. 2136
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="liver"
/tissue_lib="liver cDNA of P.M."
/note="putative"
/codon_start=1
/product="tumor necrosis factor receptor 2 related protein"
/protein_id="AAA36757.1"

CDS
1
/translation="MLPWATNAPGLAMPIVIGLGLIAASQPAVAPYASENOTCR
DOEKETEPORHICSRCPPTGVTSKRSRIRDTVCATCAENSINENHNYLTICQCR
PCDPVWGELEIAPCTSKRTQRCQPGMCAWALECHLISDPPTAELELDEV
GCDNNHVCYCKAGHONTSSPARCQPHRMOGIVEAPGTAOSDTCKNPLELP
PMSGTMLAVILPLAFLLATVESCIMKSHPSICRLGSLKRRPGEGPNVAG
SWEPRKAHPFDLIVOPILPIISGDVSPVSTGI.PAAVLEAGVPOOQSPDLTREPQLE
PGEOSVANGTNGINHYTGSMTITGNTIYNKPVLAGPGPRDDLPATPEPPPIPIEBEG
DPGPELSTPHOEDGAWHLALETHERGATPSNRPNPTTHD"

BASE COUNT 446 a 706 c 608 g 376 t

Query Match 92.5%; Score 393; DB 9; Length 2136;
Best Local Similarity 100.0%; Pred. No. 2.1e-92;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 GGATCCCTGCTCAAGAGGGGTCGGCAGAGGAGAGGCCCAATCTTGTAGTGGAGCTGG 92
DB 943 GGATCCCTGCTCAAGAGGGGTCGGCAGAGGAGAGGCCCAATCTTGTAGTGGAGCTGG 1002

ORIGIN	BASE COUNT	462 a	705 c	606 g	375 t
Query Match	92.5%	Score 393:	DB 9:	Length 2148:	
Best Local Similarity	100.0%	Pred. No. 2, 1e-92:			
Matches 393:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0:	
QY	33	GGATCGCTGCTCAAGAGCGCGTCCGACGAGGAGGAGGAGGCCAATCTCTGTAGCTGGAACCTGG	92		
DB	966	GGATCGCTGCTCAAGAGCGCGTCCGACGAGGAGGAGGAGGCCAATCTCTGTAGCTGGAACCTGG	1025		
QY	93	GAGCCTCCGAAGGCCCATCTCCATCTTCCCTGCACTTGTGTAGACACATCGTACCACTTTCT	152		
DB	1026	GAGCCTCCGAAGGCCCATCTCCATCTTCCCTGCACTTGTGTAGACACATCGTACCACTTTCT	1085		
QY	153	GGAGATGTTTCCCAAGTATCACTGAGGCTCCCGACGCCCATGTTTGAAGGAGGGGTG	212		
DB	1086	GGAGATGTTTCCCAAGTATCACTGAGGCTCCCGACGCCCATGTTTGAAGGAGGGGTG	1145		
QY	213	CCGACACAGCAGAGTCTCTGTGACCTGACCCAGGAGGCCGACGTTGGAACCCGGGAGACAG	272		
DB	1146	CCGACACAGCAGAGTCTCTGTGACCTGACCCAGGAGGCCGACGTTGGAACCCGGGAGACAG	1205		
QY	273	AGCCAGTGGGCCACGATGACCAATGGCAATGGCAATTCATGTCACGGCGGGTATGACTATCACT	332		
DB	1206	AGCCAGTGGGCCACGATGACCAATGGCAATGGCAATTCATGTCACGGCGGGTATGACTATCACT	1265		
QY	333	GGCAACATCTACATCTACAAATGAGACCACTACTTGGGGGAGACACCGGGGCTCGAGACCTC	392		
DB	1266	GGCAACATCTACATCTACAAATGAGACCACTACTTGGGGGAGACACCGGGGCTCGAGACCTC	1325		
QY	393	CCAGCTACCCCGGACGCTCCATACCCCATTTCC	425		
DB	1326	CCAGCTACCCCGGACGCTCCATACCCCATTTCC	1358		
RESULT 7	BC026262	2161 bp	mRNA	linear	PRI 08-APR-2002
LOCUS	BC026262				
DEFINITION	BC026262				
ACCESSION	BC026262				
VERSION	BC026262.1				
KEYWORDS	MGc.				
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens.				
REFERENCE	Strausberg, R.				
AUTHORS	Strausberg, R.				
TITLE	Direct Submissions				
JOURNAL	Submitted (02-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov				
COMMENT	Contact: MGC help desk Email: cgabs-r@mail.nih.gov Tissue Procurement: CLONTECH CDNA Library Preparation: CLONTECH Laboratories, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLT) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: http://www.shgc.stanford.edu Contact: (Dickson, Mark) mdickpaxl@stanford.edu Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.				

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNLT at: http://image.llnl.gov Series: IRAL Plate: 37 Row: b Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA g1: 4505038.

REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

1 (bases 1 to 178228)

Muzny,D.M., Adams,C., Adio-Oduola,B., Al-Osman,F.R., Allen,C.,
Aisbroughs,S.L., Amaralungu,H.C., Are,J.R., Ayele,M., Banks,T.,
Barberia,J., Benton,J., Blumage,K., Blankenburg,K., Bonin,D.,
Bouch,J., Bowle,S., Bileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Devilla,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Eamhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frazz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homsli,F., Howard,S., Huber,J., Huylk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratochvic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,
Lozad,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,S., Ogih,M., Okunou,G.,
Oragunye,N., Oviold,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shoshitari,N., Sisson,I.,
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Syatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,Y., Villalob,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
Williams,G., Williamson,A., Wleczek,R., Woodson,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

TITLE
JOURNAL
REFERENCE

2 (bases 1 to 178228)

Worley,K.C.

Direct Submission

Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

JOURNAL
TITLE

AUTHORS

JOURNAL

COMMENT

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: G2QW

Center clone name: CH230-362C16

Sequencing Statistics

Sequencing vector: Plasmid

Chemistry: Dye-terminator Big Dye: 100% of reads

Assembly program: Phrap, version 0.990329

Consensus quality: 116771 bases at least Q40

Consensus quality: 123909 bases at least Q30

Consensus quality: 129462 bases at least Q20

* NOTE: Estimated insert size may differ from sequencing length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_drift_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 64 contigs. The true order of the pieces
* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1	1173:	contig of 1173 bp in length
*	1174	1273: gap of unknown length
*	1274	2435: contig of 1162 bp in length
*	2436	2535: gap of unknown length
*	2536	3567: contig of 1032 bp in length
*	3568	3667: gap of unknown length
*	3668	5330: contig of 1563 bp in length
*	5231	5330: gap of unknown length
*	5331	6732: contig of 1402 bp in length
*	6733	6832: gap of unknown length
*	6833	7990: contig of 1158 bp in length
*	7991	8090: gap of unknown length
*	8091	9641: contig of 1551 bp in length
*	9642	9741: gap of unknown length
*	9742	10776: contig of 1035 bp in length
*	10777	10876: gap of unknown length
*	10877	11905: contig of 1029 bp in length
*	11906	12005: gap of unknown length
*	12006	13320: contig of 1315 bp in length
*	13321	13420: gap of unknown length
*	13421	14678: contig of 1258 bp in length
*	14679	14778: gap of unknown length
*	14779	15851: contig of 1073 bp in length
*	15852	15951: gap of unknown length
*	15952	17593: contig of 1642 bp in length
*	17594	17693: gap of unknown length
*	17694	18911: contig of 1218 bp in length
*	18912	19011: gap of unknown length
*	19012	20145: contig of 1134 bp in length
*	20146	20245: gap of unknown length
*	20246	21356: contig of 1311 bp in length
*	21357	21656: gap of unknown length
*	21657	22715: contig of 1059 bp in length
*	22716	22815: gap of unknown length
*	22816	23949: contig of 1134 bp in length
*	23950	24049: gap of unknown length
*	24050	25331: contig of 1282 bp in length
*	25332	25431: gap of unknown length
*	25432	26177: contig of 1186 bp in length
*	26178	26717: gap of unknown length
*	26718	28435: contig of 1718 bp in length
*	28436	28535: gap of unknown length
*	28536	30660: contig of 2125 bp in length
*	30661	30760: gap of unknown length
*	30761	32263: contig of 1503 bp in length
*	32264	32363: gap of unknown length
*	32364	34165: contig of 1802 bp in length
*	34166	34265: gap of unknown length
*	34266	35302: contig of 1037 bp in length
*	35303	35402: gap of unknown length
*	35403	37282: contig of 1880 bp in length
*	37283	37382: gap of unknown length
*	37383	39160: contig of 1778 bp in length
*	39161	39260: gap of unknown length
*	39261	40880: contig of 1620 bp in length
*	40881	40980: gap of unknown length
*	40981	42436: contig of 1456 bp in length
*	42437	42536: gap of unknown length
*	42537	44381: contig of 1845 bp in length
*	44382	44481: gap of unknown length
*	44482	46391: contig of 1910 bp in length
*	46392	46491: gap of unknown length
*	46492	48409: contig of 1918 bp in length
*	48410	48509: gap of unknown length
*	48510	49931: contig of 1422 bp in length
*	49932	50031: gap of unknown length
*	50032	51371: contig of 1340 bp in length
*	51372	51471: gap of unknown length

```

* 51472 53362: contig of 1891 bp in length
* 53363 53462: gap of unknown length
* 53463 55170: contig of 1708 bp in length
* 55171 55270: gap of unknown length
* 55271 56697: contig of 1427 bp in length
* 56698 56797: gap of unknown length
* 56798 58037: contig of 1240 bp in length
* 58038 58137: gap of unknown length
* 58138 60361: contig of 2224 bp in length
* 60362 60461: gap of unknown length
* 60462 62335: contig of 1874 bp in length
* 62336 62435: gap of unknown length
* 62436 64906: contig of 2471 bp in length
* 64907 65006: gap of unknown length
* 65007 67475: contig of 2469 bp in length
* 67476 67575: gap of unknown length
* 67576 69653: contig of 2078 bp in length
* 69654 73830: gap of unknown length
* 73831 73930: contig of 4077 bp in length
* 73931 77047: gap of unknown length
* 77048 77147: contig of 3117 bp in length
* 77148 80438: gap of unknown length
* 80439 80538: gap of unknown length
* 80539 83532: contig of 2994 bp in length
* 83533 83632: gap of unknown length
* 83633 85754: gap of 2122 bp in length
* 85755 85854: gap of unknown length
* 85855 89203: contig of 3349 bp in length
* 89204 92865: gap of unknown length
* 92866 92965: gap of 3562 bp in length
* 92966 95763: gap of 2798 bp in length
* 95764 95863: gap of unknown length
* 95864 99698: contig of 3835 bp in length
* 99699 99798: gap of unknown length
* 104352 104352: contig of 4554 bp in length
* 104353 104452: gap of unknown length
* 104453 108165: contig of 3713 bp in length
* 108166 108265: gap of unknown length
* 108266 111945: contig of 3680 bp in length
* 111946 112045: gap of unknown length
* 112046 116301: contig of 4256 bp in length
* 116302 116401: gap of unknown length
* 116402 120825: contig of 4424 bp in length
* 120826 120925: gap of unknown length

```

```

Query Match 27.7%; Score 117.6; DB 2; Length 178228;
Best Local Similarity 72.2%; Pred. No. 2.6e-20;
Matches 153; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

```

```

OY 89 CTGGAGCCCTCCGAGGCCATCTTCTGACTTGGTACGACGACCTGACCAT 148
    ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 152659 CTCCGAGCCCTCCGAGGCCATCTTCTGACTTGGTACGACGACCAT 152600
OY 149 TTCTGAGATGTTCTCCCACTATTCACCTGGCTCCCGACGACGACGAG 208
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 152599 GTCTGAGATGTTCTCCCACTATTCACCTGGCTCCCGACGACGAG 152540
OY 209 GGTGCGCCCAACAGACAGTCTCTGACCTGACGAGGAGCCGACGAGGAG 268
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 152539 GGTGCTACACAGACATGATGTCCTCCGATCCAGGACGAGGAGGAG 152480
OY 269 GCAGAGCCAGAGTGGCCCGACGATCAATGCGCA 300
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 152479 ACATGGCCAGTGGCCCGACGATGAGTGGCGCA 152448

```

```

RESULT 12
LOCUS AC125909 187998 bp DNA linear HTG 23-JUL-2002
DEFINITION Rattus norvegicus clone CH230-100N7, *** SEQUENCING IN PROGRESS
***, 64 unordered pieces.

```

ACCESSION AC125909
VERSION AC125909.1 GI:21671507
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 187998)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alshrocks,S.L., Amaralunge,H.C., Are,J.R., Ayale,M., Banks,T.,
Barbaria,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Eamhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Kovach,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louieged,H.,
Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabadi,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokoko,S., Ogih,M., Okunou,G.,
Orangunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Qullies,M., Ren,Y.G.,
Rives,M., Rojas,A., Kojubokan,I., Rolfe,M., Ruiz,S., Saverly,G.,
Scherer,S., Scott,G., Shen,H., Shoshart,I., Sisson,I.,
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,C., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczko,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G., and Gibbs,R.

TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 187998)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (02-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 187998)
Worley,K.C.
REFERENCE Direct Submission
AUTHORS Submitted (23-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
COMMENT ----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GVS4
Center clone name: CH230-100N7
----- Summary Statistics
Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap: version 0.990329
Consensus quality: 133860 bases at least Q40
Consensus quality: 142697 bases at least Q30
Consensus quality: 148841 bases at least Q20

NOTE: Estimated insert size may differ from sequence length
(see http://www.bjsgc.bcm.tmc.edu/docs/genbank/draft_data.html).
NOTE: This is a 'working draft' sequence. It currently
consists of 64 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 1031: contig of 1031 bp in length
1032 1131: gap of unknown length
1132 2334: contig of 1203 bp in length
2335 2434: gap of unknown length
2435 3488: contig of 1054 bp in length
3489 3588: gap of unknown length
3589 4860: contig of 1272 bp in length
4861 4960: gap of unknown length
4961 6021: contig of 1061 bp in length
6022 6121: gap of unknown length
6122 7174: contig of 1053 bp in length
7175 7274: gap of unknown length
7275 8532: contig of 1258 bp in length
8533 8632: gap of unknown length
8633 9711: contig of 1079 bp in length
9712 9811: gap of unknown length
9811 11308: contig of 1497 bp in length
11309 11408: gap of unknown length
11409 12604: contig of 1196 bp in length
12605 12704: gap of unknown length
12705 14114: contig of 1410 bp in length
14115 14214: gap of unknown length
14215 16706: contig of 2492 bp in length
16707 16806: gap of unknown length
16807 17875: contig of 1069 bp in length
17876 17975: gap of unknown length
17976 19654: contig of 1679 bp in length
19655 19754: gap of unknown length
19755 21683: contig of 1929 bp in length
21684 21783: gap of unknown length
21784 22871: contig of 1088 bp in length
22872 22971: gap of unknown length
22972 24666: contig of 1695 bp in length
24667 24766: gap of unknown length
24767 26943: contig of 2177 bp in length
26944 27043: gap of unknown length
27044 29057: contig of 2014 bp in length
29058 29157: gap of unknown length
29158 30973: contig of 1816 bp in length
30974 31073: gap of unknown length
31074 32838: contig of 1765 bp in length
32839 32938: gap of unknown length
32939 34616: contig of 1678 bp in length
34617 34716: gap of unknown length
34717 37198: contig of 2482 bp in length
37199 37298: gap of unknown length
37299 39477: contig of 2179 bp in length
39478 39577: gap of unknown length
39578 41815: contig of 2238 bp in length
41816 41915: gap of unknown length
41916 44488: contig of 2573 bp in length
44489 44588: gap of unknown length
44589 46062: contig of 1474 bp in length
46063 46162: gap of unknown length
46163 48328: contig of 2166 bp in length
48329 48428: gap of unknown length
48429 51395: contig of 2967 bp in length

51396 51495: gap of unknown length
51496 54159: contig of 2664 bp in length
54160 54259: gap of unknown length
54260 56383: contig of 2124 bp in length
56384 56483: gap of unknown length
56484 58779: contig of 2296 bp in length
58780 58879: gap of unknown length
58880 60911: contig of 2032 bp in length
60912 61011: gap of unknown length
61012 63907: contig of 2896 bp in length
63908 64007: gap of unknown length
64008 66202: contig of 2195 bp in length
66203 66302: gap of unknown length
66303 69738: contig of 3436 bp in length
69739 69838: gap of unknown length
69839 72410: contig of 2572 bp in length
72411 72510: gap of unknown length
72511 76219: contig of 3709 bp in length
76220 76319: gap of unknown length
76320 79029: contig of 2710 bp in length
79030 79129: gap of unknown length
79130 83311: contig of 4182 bp in length
83312 83411: gap of unknown length
83411 86394: contig of 2983 bp in length
86395 86494: gap of unknown length
86495 88582: contig of 2088 bp in length
88583 88682: gap of unknown length
88683 90972: contig of 2290 bp in length
90973 91072: gap of unknown length
91073 93875: contig of 2803 bp in length
93876 93975: gap of unknown length
93976 98771: contig of 4796 bp in length
98772 98871: gap of unknown length
98872 101999: contig of 3128 bp in length
102000 102099: gap of unknown length
102100 105352: contig of 3253 bp in length
105353 105452: gap of unknown length
105453 109873: contig of 4421 bp in length
109874 109973: gap of unknown length
109974 112881: contig of 2908 bp in length
112882 112981: gap of unknown length
112982 117279: contig of 4298 bp in length
117280 117379: gap of unknown length
117380 121634: contig of 4255 bp in length
121635 121734: gap of unknown length
121735 125920: contig of 4186 bp in length
125921 126020: gap of unknown length
126021 129909: contig of 3889 bp in length
130009 130010: gap of unknown length
130010 135724: contig of 5714 bp in length
135723 135823: gap of unknown length

Query Match 19.4% Score 82.4; DB 2; Length 187998;
Best Local Similarity 81.9% Pred. No. 4.3e-11;
Matches 95; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 310 CCGGGGGTCTATGCTCTACCTGGACACATCTACATCTACAATGACAGCTAGTGGGG 369
DB 154396 CTGGAGGCTCTGTGACTCTGACCGGCAATATCTACATATACAAAGGCGCAATGCTGGGG 154455
QY 370 GACACACGGGACCTGAGAGACCTCCAGCTACCCCGAAGCTCCATACCCCATTTCC 425
DB 154456 GAACACGGGGCCCTGGAGACCCCTCAGCTCCCGCTGAGCTCCATACCCGACCTCC 154511

RESULT 13
AF429315/c 125020 bp DNA linear PRI 18-JAN-2002
LOCUS Homo sapiens junctophilin 3 (JPH3) gene, partial cds.
DEFINITION AF429315
ACCESSION AF429315
VERSION AF429315.1 GI:17646244
KEYWORDS Homo sapiens.
SOURCE

ORGANISM	Homo sapiens
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 125020)
TITLE	Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S., Ingersoll-Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A., Potter,N.T., Ross,C.A. and Margolis,R.L. A repeat expansion in the gene encoding juncctophilin-3 is associated with Huntington disease-like 2
JOURNAL	Nat. Genet. 29 (4), 377-378 (2001)
MEDLINE	2 (bases 1 to 125020)
PUBMED	21583737 11694876
REFERENCE	Holmes,S.E., Ingersoll-Ashworth,R.G., Ross,C.A. and Margolis,R.L. Direct Submission Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA
JOURNAL	Location/Qualifiers
FEATURES	1..125020
SOURCE	/organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="16" /map="16q24.3; between D16S520 and Wt-12410" /note="Isolated from a patient with Huntington's disease-like 2 (HD12)" complement(35581..35746) /rpt_type=tandem /rpt_unit=ctg complement(<35507..>36887) /gene="JPH3" /note="JPH3" complement(<35507..>36887) /gene="JPH3" /product="Juncctophilin 3" complement(<36507..36887) /gene="JPH3" /note="component of the junctional complex between plasma membrane and endoplasmic reticulum" codon_start=1 /product="Juncctophilin 3" /protein_id="AA140941.1" /db_xref="GI:17646245" /translation="MSSGGRNEDPDGSGYCGGMDKAHGVGTGKGGCEYSGMS HGFEVLATWPSGNTYQGTWAGKRGITGLSEKGMVYKEMTHGFRGGRVREGAG NGARYESTGWSGLDQGYGTETYSDG"
CDS	BASE COUNT 29056 a 32731 c 30696 g 28283 t 4254 others ORIGIN
Query Match	9.4%; Score 40; DB 9; Length 125020; Best Local Similarity 11.8%; Pred. No. 5.6;
Matches	36; Conservative 166; Mismatches 97; Indels 5; Gaps 2;
QY	11 GCACATGTGCCCCAGAGATGTCAGATTCGTGCTCAAGAGCGGTCCGACGAGGAGGAGACC 70
Db	51038 SGCCTSKTSCSISGRISSTCTCGGCMCGKMSGCCASNGCK-STYKMGSMCRGRRGWS 50980
QY	71 CAATCCTGTAGCTGGAAGCTGGAGACCTCCGAAAGCCCATCATCACTTCCCTGACTTGGT 130
Db	50979 CMYWSMKSYSMAYSKSSARGRGMCMCYTSCMKYYIYSMBKSSHDBSGYKRYBYWT 50920
QY	131 ACAGCCACTGCTACCCACTATTTCTGGAGAGTTTCCCCAGTATTCCTAGCTGGCTCCCGCAGC 190
Db	50919 CMMKSMCMWSVDSHMRDMYDKMKMKYSSGTRGHBSAWMKCMSRSDSGSG--- 50864
QY	191 CCCACTTTTGGAGGAGGGGTGGCCGACAGAGAGTCTCTGAGACCTGACAGGAGGAGCC 250
Db	50863 MTHMKSYSKSSSKIMYISGKMSYVBYDCKMSWYTSBSKSYMBRWSGYSYRRYKCAW 50804
QY	251 GCAGTTTGAACCCGGGACGACAGGCTGAGGCCACAGGATCAATGAGCATCATGTCTAC 310
Db	50803 RRYBHRSSRYSWYGGKSCSWRDVSYRRRSRYHTSMSSMRMCRGARGARAKCAGAKMGSV 50744
QY	311 CGGC 314

```

RESULT 14
AF429315      125020 bp   DNA       linear    PRI 18-JAN-2002
LOCUS        AF429315
DEFINITION   Homo sapiens junctophilin 3 (JPH3) gene, partial cds.
ACCESSION    AF429315
VERSION      AF429315.1
KEYWORDS     GI:17646244
SOURCE       Homo sap lens.
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 125020)
AUTHORS      Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,
              Ingersoll-Ashworth,R.G., Fleisher,A., Stevanid,G., Brice,A.,
              Potter,N.T., Ross,C.A. and Margolis,R.L.
              A repeat expansion in the gene encoding junctophilin-3 is
              associated with Huntington disease-like 2
              Nat. Genet. 29 (4), 377-378 (2001)
TITLE        JOURNAL
MEDLINE      21583737
PUBMED       11694876
REFERENCE    2 (bases 1 to 125020)
AUTHORS      Holmes,S.E., Ingersoll-Ashworth,R.G., Ross,C.A. and Margolis,R.L.
              Direct Submission
              Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical
              Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA
JOURNALNL
FEATURES
SOURCE
location/Cuallifiers
1..125020
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/map="16q24.3; between D16S520 and WI-12410"
/note="Isolated from a patient with Huntington's
Disease-Like 2 (HDL2)"
complement(3581..35746)
/rpt_type=tandem
/rpt_unit=ctg
complement(<36507..>36887)
/gene="JPH3"
/note="JPH3"
complement(<36507..>36887)
/gene="JPH3"
/product="junctophilin 3"
complement(<36507..>36887)
/gene="JPH3"
/note="Component of the junctional complex between plasma
membrane and endoplasmic reticulum"
/codon_start=1
/product="junctophilin 3"
/protein_id="AALA0941"
/db_xref="GI:17646245"
/translaction="MSGSGRNPDDGSGYCGGWEDGAHGHCVTCPKGGCETYSMS
HGFEVLGVTPSGNTYQGTWAQGKRIGTLESKGKWYKEWHGFRGVRVRCAG
NGARVEGTSWNGLDGDGYGTETYSBG"
BASE COUNT  29056 a 32731 c 30696 g 28283 t 4254 others
ORIGIN
Query Match          9.0%, Score 38.4, DB 9, Length 125020;
Best Local Similarity 9.5%; Pzed. No.15;
Matches 39; Conservative 196; Mismatched 172; Indels 5; Gaps 1;
DB 17350 ACCCAATCCCTAGACTGGAAGCTGGAGAGCCCATTCATCTTCCCTGACTT 127
|::: : :::::::::::::: | :::::::::::::: ::::::::::::::
17350 TGCYCWMSSKSKCYKYMMMRKRRKKMKWKMSGCMWSMMGSAASRSSCYKCYKSRCSMSS 17409
|::: : :::::::::::::: | ::::::::::::::
128 GTTACAGCCACTGCTACCACCAATTCTTGAGAAGTTTCCCCAGATATCCACTGGGCTCCCCGC 187

```


GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:05:15 ; Search time 73.9741 Seconds
(without alignments)
12938.300 Million cell updates/sec

Title: US-09-917-372-2_COPY_853_1277

Perfect score: 425
Sequence: 1 atccatgagccactgcc.....aacctcatcaccatccccc 425

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N.Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	393	92.5	1594	24	ABQ55003
2	393	92.5	2136	24	ABN95637
3	393	92.5	2136	24	ABK64502
4	393	92.5	2136	24	ABL64078
5	393	92.5	2136	24	ABL64384
6	253.8	59.7	281	24	ABL82623
7	251.2	59.1	289	24	ABL82688
8	36.2	8.5	450	23	ABV38197
9	35.2	8.3	333	20	AAV72224

10	34.2	8.0	268	13	AAQ25135
11	33.6	7.9	33675	24	AAD31202
12	33.2	7.8	955	22	AA501024
13	33.2	7.8	1007	22	AA501022
14	33	7.8	1057	21	AAZ61623
15	33	7.8	1057	22	AAQ95556
16	33	7.8	1057	24	ABL34708
17	33	7.8	1057	24	AA784865
18	33	7.8	1839	20	AAZ28415
19	33	7.8	1890	22	AAQ97999
20	33	7.8	1880	24	ABL34951
21	33	7.8	1894	20	AAZ28424
22	33	7.8	2004	21	AAZ61757
23	33	7.8	2004	22	AAQ95960
24	33	7.8	2004	24	ABL34842
25	33	7.8	13085	23	ABL03580
26	33	7.8	16818	22	ABA13470
27	33	7.8	16818	22	ABA16242
28	33	7.8	16818	22	ABA18914
29	32.8	7.7	186	16	AAZ26531
30	32.8	7.7	376	19	AAV59700
31	32.6	7.7	1527	23	AA568546
32	32.6	7.7	1620	23	AA568547
33	32.6	7.7	2178	23	AA586106
34	32.6	7.7	3112	23	ABL05387
35	32.6	7.7	5721	23	ABL05386
36	32.4	7.6	649	24	ABQ72658
37	32.4	7.6	1056	21	AA551683
38	32.4	7.6	1276	24	ABQ72500
39	32.4	7.6	4677	23	AA588638
40	32.4	7.6	4677	23	AA589696
41	32.4	7.6	4677	23	AA592991
42	32.2	7.6	785	24	ABK39773
43	32.2	7.6	2861	12	AAQ14383
44	32.2	7.6	2861	21	AAZ21133
45	32.2	7.6	2861	21	AAA35011

ALIGNMENTS

RESULT 1	ABQ55003	ABQ55003 standard; cDNA; 1594 BP.
ID	ABQ55003	
AC	ABQ55003	
XX		
DT	22-AUG-2002 (first entry)	
XX		
DE	Human ovarian antigen HSBABJ44 cDNA, SEQ ID NO:883.	
XX		
XX	Human: ovarian antigen; ovary; ovarian; breast; cancer; tumour;	
KW	ovarian cancer; breast cancer; tumour; reproductive system disorder;	
KW	infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;	
KW	PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;	
KW	inflammatory condition; immune disorder; blood disorder;	
KW	cardiovascular disorder; respiratory disorder; neurological disorder;	
KW	gastrointestinal disorder; urinary system disorder; drug screening;	
KW	gene therapy; chromosome mapping; forensic analysis;	
KW	antibody preparation; cytostatic; immunomodulatory; neuroprotective;	
KW	antiinflammatory; gynaecological; reproductive; chromosome 12p13;	
KW	gene; ss.	
OS	Homo sapiens.	
XX		
PN	WO200200677-A1.	
XX		
PD	03-JAN-2002.	
XX		
PF	07-JUN-2001: 2001MO-US18569.	
XX		
PR	07-JUN-2000: 2000US-209467P.	
XX		

DR1. Synthetic.
Oryza sativa gcpe
Sugarcane plant ge
Sugarcane plant ge
cDNA encoding rat
Skin cell cDNA, SE
Skin cell cDNA, SE
Scapulariopsis bre
S. brevicaulis bet
Skin cell cDNA, SE
Rat cDNA isolated
S. brevicaulis bet
cDNA encoding rat
Skin cell cDNA, SE
Rat cDNA isolated
Drosophila melanog
Human nervous syst
Human nervous syst
Human nervous syst
Human gene signalu
Human secreted pro
DNA encoding novel
DNA encoding novel
DNA encoding novel
Drosophila melanog
Drosophila melanog
Human MDR1 encodin
Zea mays DNA fragm
Human MDR1 encodin
Human MDR1 encodin
DNA encoding novel
DNA encoding novel
DNA encoding novel
cDNA encoding lung
Clone 1 encoding I
Human low adenostin
Human adenostin re

PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA:
XX
XX WPI: 2002-147878/19.
DR P-PSDB: ABP41926.
XX
PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g.
PT ovarian cancer), immune disorders, cardiovascular disorders and
PT neurological diseases -
PS
PS Claim 1; SEQ ID NO 883; 2922pp; English.
XX
XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC ABP43228) and to cDNAs encoding them (ABO54131-ABO56305), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human ovarian antigen
CC polynucleotides, antibodies against human ovarian antigens, and the use
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC treating, prognosing or preventing ovarian cancer and/or breast-related
CC disorders. Such conditions include ovarian cancer and breast cancer, and
CC metastatic tumours of ovarian or breast origin, reproductive system
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents cDNA encoding a human ovarian antigen of the
CC invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1594 BP; 391 A; 488 C; 434 G; 280 T; 1 other:
Query Match 92.5%; Score 393; DB 24; Length 1594;
Best Local Similarity 100.0%; Pred. No. 2.8e-100;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 695 GGCAACATCTACATCTACATGACGAGTACTGGGGGACCCAGGGGCTCTGGAGACTC 754
QY 393 CCAGCTACCCCCGAGACTCCATACCCATTCCC 425
DB 755 CCAGCTACCCCCGAGACTCCATACCCATTCCC 787
RESULT 2
ABN95637
ID ABN95637 standard; DNA: 2136 BP.
XX
XX ABN95637;
AC
DE 13-AUG-2002 (first entry)
XX
XX Gene #2135 used to diagnose liver cancer.
DE
XX Gene: liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumour; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
XX Homo sapiens.
OS
XX WO200229103-A2.
PN
XX 11-APR-2002.
PD
XX 02-OCT-2001; 2001WO-US30589.
PF
XX 02-OCT-2000; 2000US-237054P.
PR
XX (GENE-) GENE LOGIC INC.
PA
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX WPI: 2002-426119/45.
DR
XX
XX Diagnosing and detecting the progression of liver cancer,
PT hepatocellular carcinoma or metastatic liver tumor in a patient,
PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample -
XX
XX Claim 1; SEQ ID NO 2135; 298pp; English.
XX
XX The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2136 BP; 446 A; 706 C; 608 G; 376 T; 0 other:
Query Match 92.5%; Score 393; DB 24; Length 2136;
Best Local Similarity 100.0%; Pred. No. 3.1e-100;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 153 GGAGATGTTTCCCACTATCCACTGGGCTCCCGCAGCCCCAGTTTGGAGGAGGGGTG 212
 |||||||
 DB 1063 GGAGATGTTTCCCACTATCCACTGGGCTCCCGCAGCCCCAGTTTGGAGGAGGGGTG 1122
 OY 213 CCGCAACAGCAGAGTCTCTGAGCCTGAGCAGGAGGAGCCAGTTGGACCCGGGAGCAG 272
 |||||||
 DB 1123 CCGCAACAGCAGAGTCTCTGAGCCTGAGCAGGAGGAGCCAGTTGGACCCGGGAGCAG 1182
 OY 273 AGCCAGGTGGCCGAGGCTACCAATGAGCATTTGTCACCGGGGGTCTATGACTATCACT 332
 |||||||
 DB 1183 AGCCAGGTGGCCGAGGCTACCAATGAGCATTTGTCACCGGGGGTCTATGACTATCACT 1242
 OY 333 GGCACATGTCATCTACATGAGACGAGTACGAGGAGGAGCAGCCGGGTCTGGAGACCTC 392
 |||||||
 DB 1243 GGCACATGTCATCTACATGAGACGAGTACGAGGAGGAGCAGCCGGGTCTGGAGACCTC 1302
 OY 393 CCAGCTACCCCGCAACCTCCATACCCCATTTCC 425
 |||||||
 DB 1303 CCAGCTACCCCGCAACCTCCATACCCCATTTCC 1335

RESULT 3

ABR64502
 ID ABR64502 standard; DNA; 2136 BP.

AC ABR64502;

DT 18-JUN-2002 (first entry)

XX Human benign prostatic hyperplasia gene #397.

XX Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.

OS Homo sapiens.

PN WO200212440-A2.

PD 14-FEB-2002.

PF 07-AUG-2001; 2001WO-US24708.

PR 07-AUG-2000; 2000US-223323P.

PR 05-JUN-2001; 2001US-0873319.

XX (GENE-) GENE LOGIC INC.

PA (MISB) JAPAN TOBRACO INC.

PI Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;

DR WPI; 2002-257476/30.

PT Identifying drugs for and diagnosing benign prostatic hyperplasia, by

PT detecting expression levels of one or more genes in prostate cells from

PT patient that are differentially regulated compared to normal prostate

PS cells

PS Disclosure; Page 229-230: 444pp: English.

XX The invention relates to a method of diagnosing (I) the onset or
 CC progression of benign prostatic hyperplasia (BPH), or screening (II) for
 CC or identifying an agent that modulates the onset or progression of BPH.
 CC The method is based on changes in gene expression in BPH tissue isolated
 CC from patients exhibiting different clinical states of prostate
 CC hyperplasia as compared to normal prostate tissue. (I) comprises
 CC detecting the expression levels of one or more genes in prostate cells
 CC from the subject that are differentially regulated compared to normal
 CC prostate cells. (II) comprises preparing a first gene expression profile
 CC of BPH cells or BPH-like cell population, exposing the cells to the
 CC agent, preparing a second gene expression profile of the agent exposed
 CC cells, and comparing the first and second gene expression profiles.
 CC (I) is useful for diagnosing the onset or progression of BPH. (II) is
 CC useful for identifying an agent that modulates the onset or progression

CC of BPH. The methods are useful to present information identifying
 CC the expression level in a tissue or cells, by comparing the expression
 CC level of genes given in a tissue or cells, in the tissue or cells to the
 CC level of expression of gene in the database, and displaying the
 CC expression levels of at least one gene in the tissue or cell sample
 CC compared to the expression level in BPH. Agents using (II) are useful for
 CC treating BPH or prostate cancer. ABR64106-ABR64860 represent human
 CC benign prostatic hyperplasia gene sequences of the invention.

XX Sequence 2136 BP; 446 A; 706 C; 608 G; 376 T; 0 other:

Query Match 92.5%; Score 393; DB 24; Length 2136;

Best Local Similarity 100.0%; Pred. No. 3,1e-100;

Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 GGATCGTCTCTCAAGAGGCGTCCGAGGAGAGGAGGCCAATCTGTAGCTGGAGACTGG 92

DB 943 GGATCGTCTCTCAAGAGGCGTCCGAGGAGAGGAGGCCAATCTGTAGCTGGAGACTGG 1002

OY 93 GAGCCCTCGAAGGCGCATCATCTTCCCTGACTGTGTACAGGCACTGCTACCATTTCT 152

DB 1003 GAGCCCTCGAAGGCGCATCATCTTCCCTGACTGTGTACAGGCACTGCTACCATTTCT 1062

OY 153 GGAGATGTTTCCCACTATCCACTGGGCTCCCGCAGCCCCAGTTTGGAGGAGGGGTG 212

DB 1063 GGAGATGTTTCCCACTATCCACTGGGCTCCCGCAGCCCCAGTTTGGAGGAGGGGTG 1122

OY 213 CCGCAACAGCAGAGTCTCTGAGCCTGAGCAGGAGGAGCCAGTTGGACCCGGGAGCAG 272

DB 1123 CCGCAACAGCAGAGTCTCTGAGCCTGAGCAGGAGGAGCCAGTTGGACCCGGGAGCAG 1182

OY 273 AGCCAGGTGGCCGAGGCTACCAATGAGCATTTGTCACCGGGGGTCTATGACTATCACT 332

DB 1183 AGCCAGGTGGCCGAGGCTACCAATGAGCATTTGTCACCGGGGGTCTATGACTATCACT 1242

OY 333 GGCACATGTCATCTACATGAGACGAGTACGAGGAGGAGCAGCCGGGTCTGGAGACCTC 392

DB 1243 GGCACATGTCATCTACATGAGACGAGTACGAGGAGGAGCAGCCGGGTCTGGAGACCTC 1302

OY 393 CCAGCTACCCCGCAACCTCCATACCCCATTTCC 425

DB 1303 CCAGCTACCCCGCAACCTCCATACCCCATTTCC 1335

RESULT 4
ABR64078
ID ABR64078 standard; DNA; 2136 BP.

XX ABR64078;

DT 15-MAY-2002 (first entry)

DE Breast cancer related gene sequence SEQ ID NO:2415.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;

XX stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;

XX cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;

XX gene; ds.

XX Homo sapiens.

PN WO200194629-A2.

PD 13-DEC-2001.

PF 30-MAY-2001; 2001WO-US10838.

PR 05-JUN-2000; 2000US-209473P.

PR 18-SEP-2000; 2000US-233133P.

PR 20-SEP-2000; 2000US-234009P.

PR 20-SEP-2000; 2000US-234034P.

PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 PA (AVALON PHARM.
 XX
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX
 DR WPI: 2002-188264/24.

PT Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity and
 XX determining a change in expression of a gene of a signature gene set
 XX
 PS Claim 1; SEQ ID 2415; 44pp; English.

XX The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in AB61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytosolic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.
 XX

SO Sequence 2136 BP; 446 A; 706 C; 608 G; 376 T; 0 other;

Query Match 92.5%; Score 393; DB 24; Length 2136;
 Best Local Similarity 100.0%; Pred. No. 3,1e-100; Indels 0; Gaps 0;
 Matches 393; Conservative 0; Mismatches 0;

QY 33 GGATCGCTGCTCAAGAGGCGTCCGAGGAGAGGAGACCAATCTGTAGCTGMACTGG 92
 Db 943 GGATCGCTGCTCAAGAGGCGTCCGAGGAGAGGAGACCAATCTGTAGCTGMACTGG 1002
 QY 93 GAGCTTCGGAAGGCCCATTCATCTCCCTGACTTGGTACAGCCACTGCTACCATTTCT 152
 Db 1003 GAGCTTCGGAAGGCCCATTCATCTCCCTGACTTGGTACAGCCACTGCTACCATTTCT 1062
 QY 153 GGAATGTTTCCCGAGTATCCACTGGGCTCCCGGAGCCCACTTTTGGAGGAGGGGTG 212
 Db 1063 GGAATGTTTCCCGAGTATCCACTGGGCTCCCGGAGCCCACTTTTGGAGGAGGGGTG 1122
 QY 213 CCGCAACAGCAGAGTCTCTGACCTTGACCAAGGAGCCGAGTGGAAACCCGGGAGCAG 272
 Db 1123 CCGCAACAGCAGAGTCTCTGACCTTGACCAAGGAGCCGAGTGGAAACCCGGGAGCAG 1182
 QY 273 AGCCAGGTGGCCCGCAGGTACCAATGSCATTCATGTCACCGGGGCTCTATGACTATCACT 332
 Db 1183 AGCCAGGTGGCCCGCAGGTACCAATGSCATTCATGTCACCGGGGCTCTATGACTATCACT 1242
 QY 333 GGCACATCTACATCTACATGACAGTACTGGGGGAGCACCGGGTCTGAGACTATCACT 392
 Db 1243 GGCACATCTACATCTACATGACAGTACTGGGGGAGCACCGGGTCTGAGACTATCACT 1302
 QY 393 CCAGCTACCCCGGACCTCCATACCCCATTCGCC 425
 Db 1303 CCAGCTACCCCGGACCTCCATACCCCATTCGCC 1335

RESULT 5

ABL64384
 ID ABL64384 standard; DNA; 2136 BP.

AC ABL64384;

XX 15-MAY-2002 (first entry)

XX Stomach cancer related gene sequence SEQ ID NO:2721.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; carcinous;
 KW cytosolic; gene therapy; anti-neoplastic; Wilms' tumour; adenocarcinoma;
 XX gene; ds.

XX Homo sapiens.

XX WO200194629-A2.

XX 13-DEC-2001.

XX 30-MAY-2001; 2001WO-US10838.

XX 05-JUN-2000; 2000US-209473P.
 PR 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-233133P.
 PR 18-SEP-2000; 2000US-233617P.
 PR 20-SEP-2000; 2000US-234099P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.

PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 (AVAL-) AVALON PHARM.
 PA
 XX
 XX
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX
 XX
 DR WPI: 2002-188264/24.
 XX
 XX
 PT Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set
 XX
 PS Claim 1; SEQ ID 2721; 44pp; English.
 XX
 XX
 CC The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in AB161664
 CC to AB170110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC adenocarcinoma, ovarian, kidney, prostate or pancreatic cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms tumor.
 CC
 XX
 XX
 SQ Sequence 2136 BP; 446 A; 706 C; 608 G; 376 T; 0 other;
 Query Match 92.5%; Score 393; DB 24; Length 2136;
 Best Local Similarity 100.0%; Pred. No. 3; Le-100;
 Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1123 CCGAACAAGAGAGATCTCTGTGACCTGACACAGGAGCCGCAATTGGAAACCCGGGAGACAG 1182
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 273 AGCCAGGTGGCCCGACGGTACCAATGGCATTCATGTCACCGGGGCTCATGACTATCACT 332
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1183 AGCCAGGTGGCCCGACGGTACCAATGGCATTCATGTCACCGGGGCTCATGACTATCACT 1242
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 333 GGCACATCTATCATATCAATGAGACGACTGCGGGGAGACACCGGGTCTTGGAAACCTC 392
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1243 GGCACATCTATCATATCAATGAGACGACTGCGGGGAGACACCGGGTCTTGGAAACCTC 1302
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 393 CCAGTACCCCGAAGCTCATACCCCATTCGCC 425
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1303 CCAGTACCCCGAAGCTCATACCCCATTCGCC 1335
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 RESULT 6
 ABL82623
 ID ABL82623 standard; cDNA; 281 BP.
 XX
 AC ABL82623;
 XX
 DT 17-MAY-2002 (first entry)
 XX
 DE Human ovarian cancer related cDNA clone SEQ ID NO:5601.
 XX
 KW Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PM W0200192581-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 29-MAY-2001; 2001WO-US17756.
 XX
 PR 26-MAY-2000; 2000US-207484P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Algate PA, Harlocker SL, Jones R;
 XX
 DR WPI: 2002-122075/16.
 XX
 PT Composition for therapy and diagnosis of ovarian cancer comprising
 PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
 PT polypeptide, antibody specific to polypeptide or T cell expressing
 PS polypeptide
 XX
 PS Claim 1; SEQ ID 5601; 489pp; English.
 XX
 CC The present invention describes a composition (I) comprising: carriers
 CC and immunostimulants; and a polypeptide (II) of a ovarian tumour
 CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence
 CC (S1) from the 10912 nucleotide sequences as given in ABL77023 to
 CC ABL87934, (III) encoding (II) having a sequence (S2), a T cell
 CC population of (II), or antigen presenting cells that express (II).
 CC (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to
 CC (S1) can be used for detecting ovarian cancer in a patient's biological
 CC sample preferably serum or ovarian tissue. The method comprises
 CC contacting a biological sample from a patient with (IV), detecting the
 CC amount of polynucleotide hybridising to (IV) and comparing the amount to
 CC a predetermined cutoff value and thereby detecting ovarian cancer in the
 CC patient, where the amount of polynucleotide hybridising to (IV) is
 CC detected preferably by polymerase chain reaction (PCR). (I) comprising
 CC (III) and/or (II) is useful for stimulating and/or expanding T cells
 CC specific for an ovarian tumour protein comprising contacting T cells
 CC with (III) or (II). (III) is useful in design and preparation of
 CC ribozyme molecules for inhibiting expression of the tumour polypeptides
 CC and proteins in tumour cells; and to isolate a full length gene from a
 CC suitable library e.g., a tumour cDNA library using well known
 CC techniques.

Sequence 281 BP; 58 A; 94 C; 77 G; 52 T; 0 other;
 Query Match 59.7%; Score 253.8; DB 24; Length 281;
 Best Local Similarity 97.1%; Pred. No. 1.7e-61;
 Matches 269; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 144 CCATTTCCTGGAGATGTTTCCCAATGATCCAGTGGGTCCCGCCAGGCC-CAGTTTGA 202
 |||
 Db 5 CCACTTCTGGAGATGTTTCCCAATGATCCAGTGGGTCCCGCCAGGCCAGCTTTTGA 64

QY 203 GGCAGGGGTCCCGAAGAGAGATCTCTGACCTGACGAGGAGCCGAGTTGNAAC 262
 |||
 Db 65 GGCAGGGGTCCCGAAGAGAGATCTCTGACCTGACGAGGAGCCGAGTTGNAAC 124

QY 263 CGGGAGACAGAGCCAGGTGGCCCAAGTACCATGCTTCAATGTCACCGGGGTCTAT 322
 |||
 Db 125 CGGGAGACAGACATCTAGTGGCCCAAGTACCATGCTTCAATGTCACCGGGGTCTAT 184

QY 323 GACATATCTGGAGATGTTTCCCAATGATCCAGTGGGTCCCGCCAGGCC-CAGTTTGA 382
 |||
 Db 185 GACATATCTGGAGATGTTTCCCAATGATCCAGTGGGTCCCGCCAGGCCAGCTTTTGA 244

QY 383 TGGAGACCTCCCGAAGAGATCTCTGACCTGACGAGGAGCCGAGTTGNAAC 419
 |||
 Db 245 TGGAGACCTCCCGAAGAGATCTCTGACCTGACGAGGAGCCGAGTTGNAAC 281

RESULT 7
 ABL82688/c
 ID ABL82688 standard; cDNA; 289 BP.
 XX ABL82688;
 XX 17-MAY-2002 (first entry)
 DE Human ovarian cancer related cDNA clone SEO ID NO:5666.
 XX Human ovarian cancer; ovarian tumour; cytostatic; gene; ss.
 KM Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
 XX Homo sapiens.
 OS WO200192581-A2.
 PN 06-DEC-2001.
 PD 29-MAY-2001; 2001WO-US17756.
 PF 26-MAY-2000; 2000US-207484P.
 PR (CORI-) CORIXA CORP.
 PA Algate PA, Harlocker SL, Jones R;
 XX WPI; 2002-122075/16.
 DR Composition for therapy and diagnosis of ovarian cancer comprising
 PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
 PT polypeptide, antibody specific to polypeptide or T cell expressing
 PT polypeptide
 PS Claim 1; SEQ ID 5666; 489pp; English.
 XX The present invention describes a composition (I) comprising: carriers
 CC and immunostimulants; and a polynucleotide (II) of a ovarian tumour
 CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence
 CC (SI) from the 10912 nucleotide sequences as given in ABL77023 to
 CC ABL87934, (III) encoding (II) having a sequence (S2), a T cell
 CC Population of (II), or antigen presenting cells that express (II).
 CC (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to
 CC (SI) can be used for detecting ovarian cancer in a patient's biological
 CC sample preferably serum or ovarian tissue. The method comprises
 CC contacting a biological sample from a patient with (IV), detecting the
 CC amount of polynucleotide hybridising to (IV) and comparing the amount to

a predetermined cutoff value and thereby detecting ovarian cancer in the
 CC patient, where the amount of polynucleotide hybridising to (IV) is
 CC detected preferably by polymerase chain reaction (PCR). (I) comprising
 CC (III) and/or (II) is useful for stimulating and/or expanding T cells
 CC specific for an ovarian tumour protein comprising contacting T cells
 CC with (III) or (II). (III) is useful in design and preparation of
 CC ribozyme molecules for inhibiting expression of the tumour polypeptides
 CC and proteins in tumour cells; and to isolate a full length gene from a
 CC suitable library e.g., a tumour cDNA library using well known
 CC techniques.

Sequence 289 BP; 49 A; 75 C; 93 G; 72 T; 0 other;
 Query Match 59.1%; Score 251.2; DB 24; Length 289;
 Best Local Similarity 98.5%; Pred. No. 9.1e-61;
 Matches 264; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 155 AGATGTTTCCCGAAGATGATCCAGTGGGTCCCGCCAGGCCAGTTTGAAGCCAGGGGTGCC 214
 |||
 Db 289 AGATGTTTCCCGAAGATGATCCAGTGGGTCCCGCCAGGCCAGTTTGAAGCCAGGGGTGCC 230

QY 215 GCACAGACAGA-GTCTCTGACCTGACGAGGAGCCGAGTTGNAACCCGGGAGCAGA 273
 |||
 Db 229 GCACAGACAGATGCTCTGACCTGACGAGGAGCCGAGTTGNAACCCGGGAGCAGA 170

QY 274 GCCAGGTGGCCCAAGGTACCATGCTTCAATGTCACCGGGGTCTATGACTATCAGT 333
 |||
 Db 169 GCCAGGTGGCCCAAGGTACCATGCTTCAATGTCACCGGGGTCTATGACTATCAGT 110

QY 334 GCACATCTACATCTACATGACCAAGTACCTGGGGGAGCAGCGGCTCTGAGAGACTCC 393
 |||
 Db 109 GCACATCTACATCTACATGACCAAGTACCTGGGGGAGCAGCGGCTCTGAGAGACTCC 50

QY 394 CAGCTACCCCGAAGCTCCATACCCCAT 421
 |||
 Db 49 CAGCTACCCCGAAGCTCCATACCCCAT 22

RESULT 8
 ABV38197
 ID ABV38197 standard; cDNA; 450 BP.
 XX ABV38197;
 XX 16-SEP-2002 (first entry)
 DE Human prostate expression marker cDNA 38188.
 XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 XX pharmacogenomic marker; gene; ss.
 OS Homo sapiens.
 OS WO200160860-A2.
 PN 23-AUG-2001.
 PD 20-FEB-2001; 2001WO-US05171.
 PF 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 PA (MTL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX Schlengel R, Endege WO, Monahan JE;
 PI WPI; 2001-662795/76.
 DR Novel isolated nucleic acid molecule associated with cancerous state of
 XX

PT	prostate cells and correlating with presence of prostate cancer. useful	
PT	for detecting presence of prostate cancer, stage of prostate cancer -	
PS	Claim 1; Page 7795; 11750pp; English.	
XX		
CC	The invention relates to an isolated nucleic acid molecule (I) comprising	
CC	a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the	
CC	specification or its complement. (I) is useful for:	
CC	(a) assessing whether a patient is afflicted with prostate cancer;	
CC	(b) monitoring the progression of prostate cancer in a patient;	
CC	(c) assessing the efficacy of a test compound to inhibit prostate	
CC	cancer in a patient;	
CC	(d) assessing the efficacy of a therapy for inhibiting prostate cancer	
CC	in a patient;	
CC	(e) selecting a composition for inhibiting prostate cancer in a patient;	
CC	(f) assessing the prostate cell carcinogenic potential of a compound;	
CC	(g) determining whether prostate cancer has metastasized in a patient;	
CC	(h) assessing the aggressiveness or indolence of prostate cancer in a	
CC	patient;	
CC	(I) is also useful as a pharmacodynamic or pharmacogenomic marker.	
XX		
XX	Sequence 450 BP; 90 A; 130 C; 127 G; 103 T; 0 other;	
SO		
Query Match	8.5%; Score 36.2; DB 23; Length 450;	
Best Local Similarity	51.6%; Pred. NO.1.9;	
Matches	83; Conservative 0; Mismatches 78; Indels 0; Gaps 0	
QY	30 TCAGATGCTGCTCAGAGGCGTCCGAGAGGAGAGGACCCCAATCCTTAGCTGAGAC 89	
Db	26 TCTAGATGATGCTGAGAGCGCGCCGACGTGATGATATGTCAGAAATTCGCCCTTAGC 85	
QY	90 TGGAGAGCTCCGAGAGGCCCATCATCTCCCTGACTGTGTACAGCCACTGTACCCATT 149	
Db	86 TGGTCGCGCGCGAGAGTACTTCTCTCTCTCCCTGCTGTGACATCATTTACAGACAGAA 145	
QY	150 TCTGGAGATGTTTCCCGCAGTATCCACTGTGGGCGCCCGCAGC 190	
Db	146 CATGGGGAAGCCAGCCAGCCTGCTGAGCCGGAGGAGAGC 186	
RESULT 9		
AAV72224		
ID	AAV72224 standard; DNA; 333 BP.	
XX		
AC	AAV72224;	
XX		
DT	07-SEP-1999 (first entry)	
XX		
DE	Human anti-GPIIb/IIIa antibody light chain DNA from phagemid PDG7.	
XX		
KW	Antibody: GPIIb/IIIa; human; auto-antibody; anti-idiotypic; diagnosis;	
KW	blood platelet membrane protein; predisposition; prevention; treatment;	
KW	autoimmune thrombocytopaenic purpura; ATP; fibrinogen binding; thrombi;	
KW	thrombocyte; cardiac infarction; pulmonary embolism; light chain; ds.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	misc_feature	1..60
FT		/*tag= a
FT		/note= "Framework region 1 (FR1)"
FT	misc_feature	61..99
FT		/*tag= b
FT		/note= "complementarity determining region 1 (CDR1)"
FT	misc_feature	100..144
FT		/*tag= C
FT		/note= "Framework region 2 (FR2)"
FT	misc_feature	145..165
FT		/*tag= d
FT		/note= "complementarity determining region 2 (CDR2)"
FT	misc_feature	166..261
FT		/*tag= e
FT		/note= "Framework region 3 (FR3)"
FT		

```

FT      misc_feature      262..294
FT      /tag= f
FT      /note= "complementarity determining region 3 (CDR3)"
FT      misc_feature      295..333
FT      /tag= g
FT      /note= "Framework region 4 (FR4)"
XX      PN      W09855619-A1.
XX      PD      10-DEC-1998.
XX      PE      05-JUN-1998; 98MO-EP03397.
XX      PR      08-MAY-1998; 98DE-1020663.
XX      PR      06-JUN-1997; 97DE-1023904.
XX      PR      12-DEC-1997; 97DE-1055227.
XX      PA      (ASAT-) ASAT AG APPLIED SCI & TECHNOLOGY.
XX      PI      Berchtoold P, Escher RFA;
XX      DR      WPI: 1999-105496/09.
XX      DR      P-PSDB; AAW90279.
XX      PT      Nucleic acid encoding human autoantibodies against platelet
XX      PT      glycoprotein IIb/IIIa - used for diagnosis, treatment and prevention
XX      PT      of autoimmune thrombocytopenic purpura and for modulation of
XX      PS      fibrinogen binding
XX      PS      Disclosure; Page 49-50; 93pp; German.
XX      CC      This invention describes novel nucleic acid fragments that encode human
XX      CC      auto-antibodies and anti-idiotypic antibodies against blood platelet
XX      CC      membrane protein, GPIIb/IIIa. The products of the invention are used
XX      CC      for diagnosis (including monitoring and determining predisposition),
XX      CC      prevention and treatment of autoimmune thrombocytopenic purpura (ATP)
XX      CC      and also for modulating binding of fibrinogen to thrombocytes
XX      CC      (particularly to dissolve thrombi and/or prevent their formation, e.g.
XX      CC      in cases of cardiac infarction or pulmonary embolism). Unlike murine
XX      CC      antibodies, human antibodies (hAb) do not induce adverse side effects
XX      CC      and persist for longer in vivo than small peptides.
XX      SQ      Sequence 333 BP; 59 A; 109 C; 93 G; 72 T; 0 other:
XX
XX      Query Match      8.3%; Score 35.2; DB 20; Length 333;
XX      Best Local Similarity 52.8%; Pred. No. 3.2;
XX      Matches 76; Conservative 0; Mismatches 68; Indels 0; Gaps 0
XX
QY      10 AGCCACTGCCCCAGAGATGTCAGGATCGTCTCCTCAAGAGGCGTCGACGAGAGAGGAC 69
QY      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      117 AGGCGAGGCGCCCAAACTCCATCCTTCATGTCATGATCATCAGCGGCCCTCCGCTCA 176
QY      70 CCAATCCGTGTAGCTGGAAAGCTGGAGCCCTCCGAAGAGCCATCATCTTCCCTGACTTG 129
QY      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      177 CCGATTCTCTGGCTCCAAAGTCGGGCAACCTCCCTCCCTGGCCATCCGTGGGCTCCAAATC 236
QY      130 TACAGCCACTGCTACCCATTTCTG 153
QY      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      237 TGGGATGCTGCTGACTATTACTG 260
QY
RESULT 10
ID      AAQ25135
XX      AAQ25135 standard; DNA; 268 BP.
XX
XX      AAQ25135:
XX      19-NOV-1992 (first entry)
XX
XX      DRL.
XX
XX      HLA class II; DRB; DQA; DOB; DPB; HLA-DQ; HLA-DQ; HLA-DQ; HLA-DQ;
XX      insulin-dependent diabetes; ss.

```

```
XX Synthetic.
OS
XX Key
XX misc_RNA
FT 7..21 location/Qualifiers
FT /tag= a
FT /label= Polymorphous_region_1
FT 61..81
FT /tag= b
FT /label= Polymorphous_region_2
XX
XX WO9208117-A.
XX
XX 14-MAY-1992.
XX
XX 08-OCT-1991; 91WO-US07308.
XX
XX 17-OCT-1990; 90NL-0002259.
XX
XX (EURO-) EURODIAGNOSTICS BV.
XX (BIOM) APPLIED BIOSYSTEMS INC.
XX
XX Tiliamus MG:
XX
XX WPI: 1992-183800/22.
XX P-PSDB: AAR24276.
XX
XX Method and kit for determining genotype - by comparing sequences
XX of gene family members e.g. for disease detection
XX
XX Disclosure; Fig 1: 22pp; English.
XX
XX The sequence given is an allele of the human DRB locus, DRI. This
XX allele is distinguishable from other DR alleles by the sequences
XX contained within the polymorphic regions. These alleles specific
XX sequences can differ between closely related individuals, see also
XX AA025136-45 and AAR24277-86. It can be seen by haplotype analysis that
XX the are only two distinct DR alleles and the allele specific
XX sequences identify different haplotypes.
XX Sequences from the HLA class II loci DQA, DQB and DRB can also be
XX analysed to produce similar information. The primer sequences used
XX in this analysis can be used determining a genotype by comparing the
XX nucleotide sequence of members of a gene system, where the sequences
XX to be compared are of a strongly conserved section of the genetic
XX material.
XX This method can be used to detect genetic variations associated with
XX diseases or disease symptoms, eg. immune response defects, diseases
XX associated with the HLA system, such as Hodgkin's disease, multiple
XX sclerosis and insulin-dependent diabetes. This method may also be
XX used for tissue or cell typing, eg. to determine the degree of
XX compatibility of transplants and to determine the risk of an
XX HLA-associated disease for an animal. It allows the typing of, eg.
XX HLA class I alleles which could previously only be determined by
XX serological techniques.
XX
XX Sequence 268 BP; 58 A; 60 C; 99 G; 51 T; 0 other;
XX
XX Query Match 8.0%; Score 34.2; DB 13; Length 268;
XX Best Local Similarity 54.3%; Pred. No. 5.7;
XX Matches 69; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
```

```
RESULT 11
AAD31202/C
ID AAD31202 standard; DNA; 33675 BP.
XX
XX AAD31202;
AC
XX 31-MAY-2002 (first entry)
XX
XX Oryza sativa gcpe gene.
XX
XX gcpe gene; methyl-D-erythritol phosphate; MEP; dimethylallyl diphosphate;
XX transgenic plant; isoprenoid compound; tocopherol; isopentenyl diphosphate;
XX food; feed source; transfection; single nucleotide polymorphism; SNP;
XX oxidative stress tolerance; UV tolerance; transformation; rice; plant;
XX ds.
XX
XX Oryza sativa.
XX
XX Location/Qualifiers
XX 6514..11129
XX /tag= a
XX /product= "Rice GCPE protein #1"
XX 6924..11129
XX /tag= b
XX /product= "Rice GCPE protein #2"
XX 6514..6762
XX /tag= c
XX /note= "This region corresponds to rice GCPE protein #1"
XX 6763..6923
XX /tag= d
XX /note= "This region corresponds to rice GCPE protein #1"
XX /cons.splice= (5'site:YES, 3'site:NO)
XX 6924..7019
XX /tag= e
XX 7020..7162
XX /tag= f
XX 7163..7269
XX /tag= g
XX 7270..7343
XX /tag= h
XX 7344..7444
XX /tag= i
XX 7445..7524
XX /tag= j
XX 7525..7634
XX /tag= k
XX 7635..7693
XX /tag= l
XX 7694..7813
XX /tag= m
XX 7814..7922
XX /tag= n
XX 7923..8153
XX /tag= o
XX 8154..8252
XX /tag= p
XX 8253..8369
XX /tag= q
XX 8370..8514
XX /tag= r
XX 8515..8589
XX /tag= s
XX 8590..9011
XX /tag= t
XX 9012..9071
XX /tag= u
XX 9072..9162
XX /tag= v
XX 9163..9225
XX /tag= w
XX 9226..9327
XX /tag= x
```

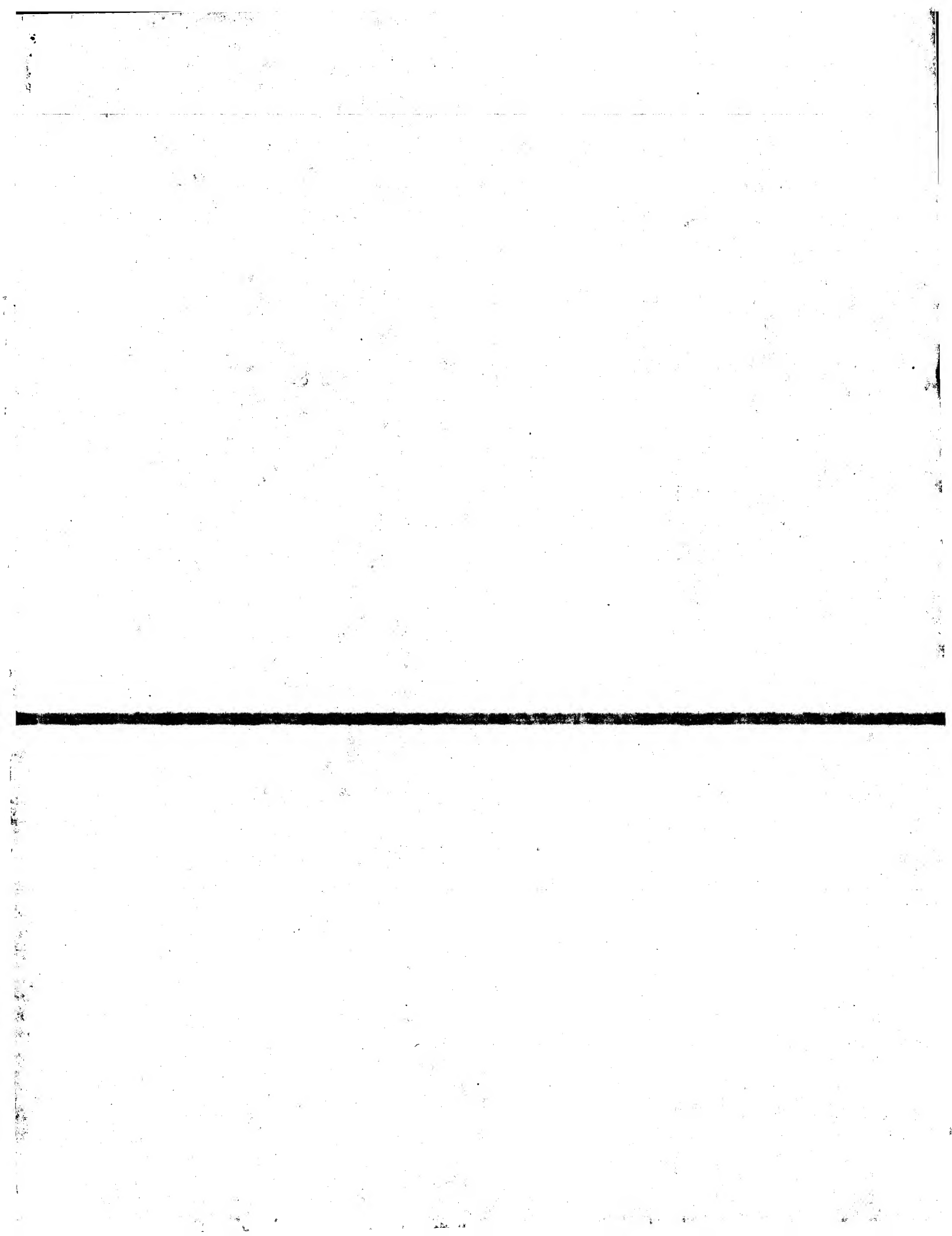
FT	exon	9328..9472	/*tag= y
FT	intron	9473..9588	/*tag= z
FT	exon	9589..9730	/*tag= aa
FT	intron	9731..9950	/*tag= ab
FT	exon	9951..10028	/*tag= ac
FT	intron	10029..10133	/*tag= ad
FT	exon	10134..10293	/*tag= ae
FT	intron	10294..10693	/*tag= af
FT	exon	10694..10798	/*tag= ag
FT	intron	10799..11027	/*tag= ah
FT	exon	11028..11129	/*tag= ai
PN			
XX		W0200212478-A2.	
XX		14-FEB-2002.	
XX		06-AUG-2001; 2001WO-US24335.	
XX		07-AUG-2000; 2000US-223483P.	
PA		(MONS) MONSANTO TECHNOLOGY LLC.	
PI	Boronat A, Campos N, Rodriguez-concepcion M, Rohmer M, Seeman M;		
PI	Valentin HE, Venkatesh TV, Venkatramesh M;		
DR		WPI: 2002-227151/28.	
DR	P-PSDB; AAE19651, AAE19652.		
XX			
PT	gcpe nucleic acid which is an essential gene of the methyl-D-erythritol		
PT	phosphate pathway, encoding a fully defined GCPE protein which is		
PT	useful for increasing levels of tocopherol substrates in plants		
XX			
PS	Claim 3; Page 97-117; 155pp; English.		
XX			
CC	The invention relates to gcpe nucleic acid molecule, an essential gene		
CC	of methyl-D-erythritol phosphate (MEP) pathway that encodes rice,		
CC	Arabidopsis thaliana or Escherichia coli GCPE protein. gcpe is useful		
CC	for producing a transgenic plant such as Brassica campestris, B. napus,		
CC	canola, castor bean, coconut, cotton, crambe, linseed, maize, mustard,		
CC	oil palm, peanut, rapeseed, rice, safflower, sesame, soybean, sunflower,		
CC	or wheat with an increased isoprenoid (tocopherol) compound level. The		
CC	expression of GCPE protein in organisms increases the level of		
CC	tocopherol substrate such as isopentenyl diphosphate and dimethylallyl		
CC	diphosphate biosynthesis. Transgenic organisms overexpressing GCPE		
CC	protein can nutritionally enhance food and feed sources. Overexpression		
CC	of GCPE protein in transgenic plant may provide tolerance to stresses		
CC	e.g., oxidative stress tolerance such as to oxygen or ozone, UV		
CC	tolerance, etc. gcpe may be used to obtain nucleic acid molecules from		
CC	the same species, and to obtain nucleic acid homologues. gcpe is also		
CC	used as or primers. The recombinant vectors are used in plant		
CC	transformation or transfection. gcpe can also act as markers capable of		
CC	detecting polymorphisms such as single nucleotide polymorphisms (SNPs).		
CC	gcpe is also used to determine the level or pattern of expression of		
CC	the protein. The present sequence is Oryza sativa gcpe gene.		
XX			
XX	Sequence 33675 BP; 9475 A; 7073 C; 7297 G; 9801 T; 29 other;		
XX			
XX	Query Match 7.9%; Score 33.6; DB 24; Length 33675;		
XX	Best Local Similarity 55.0%; Pred. No. 37;		
XX	Matches 66; Conservative 0; Mismatches 54; Indels 0; Gaps 0;		
XX			
XX	229 CTCTGACCTGACCGAGCGCGAGTTGGACCCGGGGAGCAGAGCCAGTGCCCCACG 288		

D	b2658	CTCCGGGAAGGCCGCTCCTCCTTGTTCATCGCGGCAGAAAGGGCGGTTGGACG	12599
Q	y289	GTAACAATTGCATTCATGTCACCAGGGGCTCATATCACTACATGCACATCATCT	348
D	12598	GCATCACGAGACACTCTACAACTTCGGGGAGCTCCATCTCGGACACCGCAACTTCCTCGGCT	12539

RESULT 12

AAS01024 standard; cDNA; 955 BP.

ID	AAS01024	
AC	AAS01024:	
XX		
DT	29-MAY-2001	(first entry)
DE	Sugarcane plant gene promoter cDNA isolated from clone c512.	
XX		
KW	Sugarcane promoter region; monocotyledonous plant; stem tissue;	
KW	Insecticide; herbicide; disease resistance; improved food content;	
KW	Beta-glucuronidase; GUS; starch biosynthesis; fatty acid biosynthesis;	
KW	ADP-glucose pyrophosphorylase; sucrose metabolism; clone c512; ss.	
XX		
CX	Saccharum sp.	
XX		
FH	Key	Location/Qualifiers
FT	5'UTR	1..122
FT		/+tag= a
FT	CDS	123..572
FT		/+tag= b
FT		/note= "Amino acid sequence deduced from the longest and conserved ORF"
FT	3'UTR	573..924
FT		/+tag= c
FT	polyA_site	925
XX		/+tag= d
PN	WO200118211-A1.	
PD	15-MAR-2001.	
PF	01-SEP-2000; 2000WO-AU01033.	
PR	02-SEP-1999; 99AU-0002625.	
PA	(UYQU) UNIV QUEBENS LAND.	
P1	Potter B, Birch RG;	
DR	WPI: 2001-218560/22.	
DR	P-PSDB; AAU00451.	
PT	New sugarcane plant promoters for directing expression of heterologous nucleic acids in a constitutive or tissue-specific manner in monocotyledonous plants -	
PS	Claim 1; Fig 15; 107pp; English.	
XX		
XX	The present sequence for sugarcane plant promoter cDNA isolated from clone c512 is 1 of 11 promoter regions of a transcribable DNA sequence isolated from various sugarcane cDNA clones (AAS01021-AAS01031).	
CC	Clones c51, c511 and c512 are homologous cDNA sequences. Also described are 4 promoter regions of specific transcribed DNA sequences (AAS01032-AAS01035). The nucleic acids are useful for producing transgenic plants, having an altered phenotype and for driving expression of a foreign or endogenous DNA sequence, which encode agronomic properties including insecticide, herbicide, disease resistance, stress tolerance and improved food content, or increased yields. The foreign or endogenous DNA sequence may comprise a region transcribed into an antisense RNA or ribozyme that modulates the expression of a corresponding target gene, or it may encode beta-glucuronidase (GUS), luciferase, neomycin phosphotransferase, a product conferring herbicide tolerance, a nematode affecting starch biosynthesis or modification,	



GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:09:45 ; Search time 490.898 Seconds

(without alignments)
14021.414 Million cell updates/sec

Title: US-09-917-372-2_COPY_853_1277

Perfect score: 425
Sequence: 1 atccattagacgactgcgc.....aacctccatccattccc 425

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	393	92.5	814	9 AU118203	AU118203 AU118203
2	393	92.5	868	10 BE275240	BE275240 601122062
3	393	92.5	895	14 BQ957322	BQ957322 AGENCOURT
4	393	92.5	920	12 BG036632	BG036632 602326834
5	393	92.5	954	13 B1821009	B1821009 603035664
6	393	92.5	968	14 BQ898015	BQ898015 AGENCOURT

7	388	91.3	427	10 AM630662	h85h07.Y
8	381	89.6	640	12 BG491235	602535283
9	381	89.6	714	13 B1829057	603079147
10	381	89.6	900	13 B1769498	603059047
11	381	89.6	901	12 BE740143	601359085
12	380.4	89.5	954	14 BQ930450	AGENCOURT
13	370	87.1	700	12 BG576875	602599233
14	355	83.5	749	12 BE789430	601482267
15	346.6	81.6	579	10 BE207591	bb66f10.Y
16	344	80.9	725	12 BG331666	602422673
17	343.8	80.9	606	13 B1905922	603063085
18	342	80.5	598	12 BG403820	602419660
19	326	76.7	728	13 B1765703	603046670
20	322	75.8	801	12 BF792867	602253351
21	316	74.4	458	14 BM766807	K-EST0048
22	316	74.4	501	14 BM766833	K-EST0048
23	316	74.4	537	14 BM766352	K-EST0048
24	316	74.4	572	14 BM767287	K-EST0049
25	316	74.4	666	14 BM767318	K-EST0049
26	314.4	74.0	634	14 BM746911	K-EST0021
27	311	73.2	598	14 BM723050	UT-E-EO1-
28	301.8	71.0	360	14 BM766382	K-EST0048
29	293.4	69.0	366	9 AA099514	2177h03.r
30	282	66.4	598	14 BQ305454	MRO-BT200
31	275	64.7	506	14 BM855039	K-EST0137
32	274.4	64.6	276	14 BM784961	K-EST0063
33	274.4	64.5	808	13 B1818411	603032875
34	273.4	64.3	466	14 BQ082995	K-EST0144
35	269.8	63.5	426	14 BM856797	K-EST0140
36	263.4	62.0	416	14 BQ083004	K-EST0144
37	261.8	61.6	418	14 BM855410	K-EST0138
38	253.8	59.7	281	9 AA476989	zu31d09.r
39	251.2	59.1	289	9 AA479617	zu31d09.s
40	249.2	58.6	942	13 B1818552	603033075
41	247	58.1	346	14 H59746	yr38e10.r1
42	246.6	58.0	721	9 A1336202	q144e06.x
43	245.4	57.7	530	12 BG610053	324524.MA
44	244	57.4	352	10 AM605918	RC1-HT025
45	236	55.5	706	12 BG698325	602658595

ALIGNMENTS

RESULT 1	AU118203	814 bp	mRNA	linear	EST 01-AUG-2002
LOCUS	AU118203	HEMBA1	Homo sapiens	cdna	clone HEMBA1003089 5', mRNA
DEFINITION	AU118203	sequence.			
ACCESSION	AU118203				
VERSION	AU118203.1	GI:10933231			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 814)				
AUTHORS	Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S., and Isogai,T.				
TITLE	HRI human cDNA project				
JOURNAL	unpublished (2000)				
COMMENT	Contact: Takao Isogai Genomics Laboratory Helix Research Institute 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan Tel: 81-438-52-3975 Fax: 81-438-52-3986 Email: genomics@hri.co.jp HRI human cDNA project: 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.				

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LINC2579 row: 9 column: 17
High quality sequence stop: 645.

FEATURES

source

1. 895

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6384424"
/clone_lib="NIH_MGC_40"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: prostate; Vector: pOTB7; Site:1: XhoI;
Site:2: EcoRI; CDNA made by oligo-dT priming.
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(C). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

BASE COUNT 190 a 291 c 252 g 159 t 3 others
ORIGIN

Query Match 92.5%; Score 393; DB 14; Length 895;
Best Local Similarity 100.0%; Pred. No. 1e-86;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

33 GGATCGCTGCTCAAGAGGCGTCCGAGAGGAGGAGCCCAATCCTGAGCTGAGACTGG 92
132 GGATCGCTGCTCAAGAGGCGTCCGAGAGGAGGAGCCCAATCCTGAGCTGAGACTGG 191
93 GAGCCTCCGAGAGGCCCATCTACTTCCCTGAGTGTGACAGGCACTGTACCACTTTCT 152
192 GAGCCTCCGAGAGGCCCATCTACTTCCCTGAGTGTGACAGGCACTGTACCACTTTCT 251
153 GGAGATGTTTCCCGACATATCCACTGGGCTCCCGAGAGCCCACTTTGGAGGAGGGTG 212
252 GGAGATGTTTCCCGACATATCCACTGGGCTCCCGAGAGCCCACTTTGGAGGAGGGTG 311
213 CGGCACACAGAGTCTCTGGACCTGACACGAGGAGCCGACAGTGTGAGACCCGGGAGCAG 272
312 CGGCACACAGAGTCTCTGGACCTGACACGAGGAGCCGACAGTGTGAGACCCGGGAGCAG 371
273 AGCCAGGTGGCCGACGATGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 332
372 AGCCAGGTGGCCGACGATGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 431
333 GGCACATCTACATCTACATGAGACGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 392
432 GGCACATCTACATCTACATGAGACGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 491
93 393 CCAGCTACCCCGAGACCTCCATACCCCACTTCCC 425
492 CCAGCTACCCCGAGACCTCCATACCCCACTTCCC 524

RESULT 4

BG036632

LOCUS 920 bp mRNA linear EST 24-JAN-2001
DEFINITION 603226834F1 NIH_MGC_91 Homo sapiens CDNA clone IMAGE:4428225 5',
mRNA sequence.

ACCESSION

BG036632

VERSION BG036632.1 GI:12432013

KEYWORDS

EST.

SOURCE

ORGANISM

Homo sapiens

human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

1 (bases 1 to 920)

TITLE NIH-MGC http://mnc.ncl.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCM/DRP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LLAM10178 row: c column: 10
High quality sequence stop: 693.

FEATURES

source

1. 920

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4428225"
/clone_lib="NIH_MGC_91"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: prostate; Vector: pCMV-Sport6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.4 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 187 a 304 c 270 g 159 t
ORIGIN

Query Match 92.5%; Score 393; DB 12; Length 920;
Best Local Similarity 100.0%; Pred. No. 1e-86;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

33 GGATCGCTGCTCAAGAGGCGTCCGAGAGGAGGAGCCCAATCCTGAGCTGAGACTGG 92
52 GGATCGCTGCTCAAGAGGCGTCCGAGAGGAGGAGCCCAATCCTGAGCTGAGACTGG 111
93 GAGCCTCCGAGAGGCCCATCTACTTCCCTGAGTGTGACAGGCACTGTACCACTTTCT 152
112 GAGCCTCCGAGAGGCCCATCTACTTCCCTGAGTGTGACAGGCACTGTACCACTTTCT 171
153 GGAGATGTTTCCCGACATATCCACTGGGCTCCCGAGAGCCCACTTTGGAGGAGGGTG 212
172 GGAGATGTTTCCCGACATATCCACTGGGCTCCCGAGAGCCCACTTTGGAGGAGGGTG 231
213 CGGCACACAGAGTCTCTGGACCTGACACGAGGAGCCGACAGTGTGAGACCCGGGAGCAG 272
312 CGGCACACAGAGTCTCTGGACCTGACACGAGGAGCCGACAGTGTGAGACCCGGGAGCAG 291
273 AGCCAGGTGGCCGACGATGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 332
372 AGCCAGGTGGCCGACGATGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 351
333 GGCACATCTACATCTACATGAGACGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 392
352 GGCACATCTACATCTACATGAGACGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 411
393 CCAGCTACCCCGAGACCTCCATACCCCACTTCCC 425
412 CCAGCTACCCCGAGACCTCCATACCCCACTTCCC 444

RESULT 5

B1821009

LOCUS 954 bp mRNA linear EST 04-OCT-2001
DEFINITION 603035664F1 NIH_MGC_115 Homo sapiens CDNA clone IMAGE:5176935 5',
mRNA sequence.

ACCESSION

B1821009

VERSION B1821009.1 GI:15932559

KEYWORDS

EST.

SOURCE

ORGANISM

Homo sapiens

human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

1 (bases 1 to 954)

TITLE NIH-MGC http://mnc.ncl.nih.gov/.

JOURNAL

AUTHORS

Db	626	COAGTACCCCGAACCCTCATACCCCACTTCC	658
RESULT 7		427 bp	mRNA
LOCUS	AM630662		linear
DEFINITION	h185h07.y1 NCI_CGAP_GUI Homo sapiens cDNA clone IMAGE:2969629 5' similar to db:L04270 TUMOR NECROSIS FACTOR RECEPTOR 2 RELATED PROTEIN PRECURSOR (HUMAN); contains element MSRI repetitive element ;, mRNA sequence.		EST 31-MAR-2000
ACCESSION	AM630662		
VERSION	AM630662.1	GI:7377452	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Mammalia: Euthera; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia: Euthera; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 427)		
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
COMMENT	Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgaps-remail.nih.gov Tissue Procurement: Chris Woskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: life Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The I.M.A.G.E. Consortium DNA Sequencing by: Washington University Genome Sequencing Center clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov/image/html/lresources.shtml Seq primer: -40RP from Glbco High quality sequence stop: 340. Location/Qualifiers 1..427 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:2969629" /clone_lib="NCI CGAP GUI" /tissue_type="2 pooled high-grade transitional cell tumors" /lab_host="DH10B" /note="Organ: genitourinary tract; Vector: pCMV-Sport6; Site_1: Salt; Site_2: NotI; Cloned unidirectionally. Primer: oligo dt. Library constructed by Life Technologies."		
BASE COUNT	89 a 145 c 118 g 75 t		
ORIGIN			
Query Match	91.3%; Score 388; DB 10; Length 427;		
Best Local Similarity	100.0%; Pred. No. 1,3e-85;		
Matches	388; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Y	38 GCTGCTCAAGAGGCGTCCGCAAGAGAGAGACCCATCTGTAGCTGAGAGCTGGAGCC	97	
Db	1 GCTGCTCAAGAGGCGTCCGCAAGAGAGAGAGACCCATCTGTAGCTGAGAGCTGGAGCC	60	
Y	98 TCCGAAGGCCCATCCATCTCTCCGACTTGGTACAGCAGCAGTCAACCATTTCTGGAGA	157	
Db	61 TCCGAAGGCCCATCCATCTCTCCGACTTGGTACAGCAGCAGTCAACCATTTCTGGAGA	120	
Y	158 TGTTCCTCCAGTATCACTTGGGCTCCCGCAAGCCAGATTTGGAGAGAGGGTCCGCA	217	
Db	121 TGTTCCTCCAGTATCACTTGGGCTCCCGCAAGCCAGATTTGGAGAGAGGGTCCGCA	180	
Y	218 ACAGCAGAGTCTCTGACCTGACAGGAGAGCCGAGTTGGAAACCCGGGAGACAGCCA	277	
Db	181 ACAGCAGAGTCTCTGACCTGACAGGAGAGCCGAGTTGGAAACCCGGGAGACAGCCA	240	
Y	278 GGTGCCCCAGGTTACCATGGATTCATGATCAACGGGGGCTATGACTATACATGGCA	337	
Db	241 GGTGCCCCAGGTTACCATGGATTCATGATCAACGGGGGCTATGACTATACATGGCA	300	

OY	338	CATCAGCATCTACAAATGAGACAGTACGGGGGAGACCAGGGGTCTCGTAGAAGCTCCAGC	397		
Db	301	CATCTACTATTCTACATATGGAACAAGTACTGGGGGAGACCAAGGGGTCTCGTAGAAGCTCCAGC	360		
OY	398	TACCCCGAACCTCCATACCCCATTTCCC	425		
Db	361	TACCCCGAACCTTCATACCCCATTTCCC	388		
RESULT 8	BG491235.	640 bp	mRNA	linear	EST 27-MAR-2001
LOCUS	BG491235				
DEFINITION	602535283FP1 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:4654076 5', mRNA sequence.				
ACCESSION	BG491235				
VERSION	BG491235.1	GI:13452747			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 640)				
TITLE	NIH-MGC http://mgc.nci.nih.gov/.				
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: rs9ab@remail.nih.gov Tissue Procurement: DCTD/DTP cDNA Library Preparation: Ling Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at: http://image.lnl.nih.gov Plate: LHCMI441 row: e column: 21 High quality sequence start: 199 High quality sequence stop: 640.				
FEATURES	Location/Qualifiers				
SOURCE	1..640				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:4654076"				
	/clone_lib="NIH_MGC_41"				
	/tissue_type="amelanotic melanoma, cell line"				
	/lab_host="DH10B (phage-resistant)"				
	/note="Organ: skin; Vector: pOTB1; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: This is a NIH-MGC library."				
BASE COUNT	134 a 205 c 159 g 142 t				
ORIGIN					
Query Match	89.6%	Score 381;	DB 12:	Length 640;	
Best Local Similarity	100.0%;	Pred. No. 8.2e-84;			
Matches 381;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
OY	33	GGATGCTGCTCAAGAGAGCGTCCGAGAGAGAGACCAATCCTGTAGCTGGAGCTGG	92		
Db	260	GGATGCTGCTCAAGAGAGCGTCCGAGAGAGAGAGACCAATCCTGTAGCTGGAGAGCTGG	319		
OY	93	GAGCCTCGGAAGGCCCATCATCTTCCCTGACTTGTAACAGCACATCTACCATTTCT	152		
Db	320	GAGCCTCGGAAGGCCCATCATCTTCCCTGACTTGTAACAGCACATCTACCATTTCT	379		
OY	153	GGAGATGTTTTCCCGAGTATTCACATGGGGTCCCGGAGGCCAGTTTGGAGGACAGGGTG	212		
Db	380	GGAGATGTTTTCCCGAGTATTCACATGGGGTCCCGGAGGCCAGTTTGGAGGACAGGGTG	439		

Db	Query Match	Score	DB	Length
141	GGAGATGTTCCCGCAATATCCACTGGGCTCCCGCAGCCCCAGTTTGGAGGCAAGGCTG	89.6%	200	900
153	GGAGATGTTCCCGCAATATCCACTGGGCTCCCGCAGCCCCAGTTTGGAGGCAAGGCTG	99.7%	212	900
201	GGAGATGTTCCCGCAATATCCACTGGGCTCCCGCAGCCCCAGTTTGGAGGCAAGGCTG		260	900
213	CCGCAACAGCAGAGTCTCTGAGCCTGGACGAGGAGCCGAGTGGAAACCGGGGAGCAG		272	900
261	CCGCAACAGCAGAGTCTCTGAGCCTGGACGAGGAGCCGAGTGGAAACCGGGGAGCAG		319	900
273	AGCCAGTGGCCCAAGGTATGCAATGATGTCACCGGGGGGTCTATGATATCACT		332	900
320	AGCCAGTGGCCCAAGGTATGCAATGATGTCACCGGGGGGTCTATGATATCACT		379	900
333	GGCAACATCTACATCTACATGACAGCAGTACTGGGGGAGCACCGGGGTCTGGAGACTC		392	900
380	GGCAACATCTACATCTACATGACAGCAGTACTGGGGGAGCACCGGGGTCTGGAGACTC		439	900
393	CCAGTATCCCCCGAAGCTCCATACCCCATTTCC		425	900
440	CCAGTATCCCCCGAAGCTCCATACCCCATTTCC		472	900
RESULT 10				
LOCUS	BI769498	900 bp	MRNA	EST 25-SEP-2001
DEFINITION	60305004/F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5208394 5'			
ACCESSION	BI769498			
VERSION	BI769498.1			
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
TITLE	1 (bases 1 to 900)			
JOURNAL	NIH-MGC http://mhc.nci.nih.gov/ .			
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Plate: L14M1522 row: n column: 11 High quality sequence stop: 854. Location/Qualifiers 1..900 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:5208394" /clone_lib="NIH_MGC_122" /lab_host="DH10B" /note="Organ: pooled lung and spleen; Vector: pCMV-Sport6; Site: 1: NotI; Site: 2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note this is a NIH_MGC Library."			
BASE COUNT	199 a 298 c 246 g 157 t			
ORIGIN				

Matches 392; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
QY 33 GGATGCTGCTCTCAAGAGGGTCCGAGAGGAGGAGCCCAATCCTGTACTGGAAGCTGG 92
Db 269 GGATGCTGCTCTCAAGAGGGTCCGAGAGGAGGAGCCCAATCCTGTACTGGAAGCTGG 328
QY 93 GAGCTCCGAGAGGCCATCATCTACTTCCCTGACTTGGTACAGCCACTGTACCCATTTC 152
Db 329 GAGCTCCGAGAGGCCATCATCTACTTCCCTGACTTGGTACAGCCACTGTACCCATTTC 388
QY 153 GAGATGTTTCCCGAGTATCCAGTGGGCTCCCGCAGCCCGAGTTTGGAGGAGGGGTG 212
Db 389 GAGATGTTTCCCGAGTATCCAGTGGGCTCCCGCAGCCCGAGTTTGGAGGAGGGGTG 448
QY 213 CGGCAACAGCAGAGTCTGTGACCTGACAGGAGGAGCCGACATTTGGAACCCGGAGGAG 272
Db 449 CGGCAACAGCAGAGTCTGTGACCTGACAGGAGGAGCCGACATTTGGAACCCGGAGGAG 507
QY 273 AGCCAGTGGCCCGACGATGACATGATGTCACCGGGGCTGTATGACTATCACT 332
Db 508 AGCCAGTGGCCCGACGATGACATGATGTCACCGGGGCTGTATGACTATCACT 567
QY 333 GGCACATCTACATCTACATGAGACCACTGAGGGGGGACCAACCGGGTCTGGAGACCTC 392
Db 568 GGCACATCTACATCTACATGAGACCACTGAGGGGGGACCAACCGGGTCTGGAGACCTC 627
QY 393 CGAGTACCCCGGACCTCCATACCCCATTTCC 425
Db 628 CGAGTACCCCGGACCTCCATACCCCATTTCC 660
```

RESULT 11

LOCUS BE740143 901 bp mRNA linear EST 15-SEP-2000
DEFINITION 601595085F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3949148 5',
mRNA sequence.

ACCESSION BE740143
VERSION BE740143.1 GI:10154135
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 901)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: ILCM813 row: 1 column: 21
High quality sequence stop: 745.

FEATURES

SOURCE

Location/Qualifiers

1..901
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3949148"
/clone_lib="NIH_MGC_9"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGCAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 181 a 297 c 267 g 156 t
ORIGIN

Query Match 89.6%; Score 381; DB 12; Length 901;
Best Local Similarity 99.7%; Pred. No. 9, 2e-84;

Matches 392; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
QY 33 GGATGCTGCTCTCAAGAGGGTCCGAGAGGAGGAGCCCAATCCTGTACTGGAAGCTGG 92
Db 55 GGATGCTGCTCTCAAGAGGGTCCGAGAGGAGGAGCCCAATCCTGTACTGGAAGCTGG 113
QY 93 GAGCTCCGAGAGGCCATCATCTACTTCCCTGACTTGGTACAGCCACTGTACCCATTTC 152
Db 114 GAGCTCCGAGAGGCCATCATCTACTTCCCTGACTTGGTACAGCCACTGTACCCATTTC 173
QY 153 GAGATGTTTCCCGAGTATCCAGTGGGCTCCCGCAGCCCGAGTTTGGAGGAGGGGTG 212
Db 174 GAGATGTTTCCCGAGTATCCAGTGGGCTCCCGCAGCCCGAGTTTGGAGGAGGGGTG 233
QY 213 CGGCAACAGCAGAGTCTGTGACCTGACAGGAGGAGCCGACATTTGGAACCCGGAGGAG 272
Db 234 CGGCAACAGCAGAGTCTGTGACCTGACAGGAGGAGCCGACATTTGGAACCCGGAGGAG 293
QY 273 AGCCAGTGGCCCGACGATGACATGATGTCACCGGGGCTGTATGACTATCACT 332
Db 294 AGCCAGTGGCCCGACGATGACATGATGTCACCGGGGCTGTATGACTATCACT 353
QY 333 GGCACATCTACATCTACATGAGACCACTGAGGGGGGACCAACCGGGTCTGGAGACCTC 392
Db 354 GGCACATCTACATCTACATGAGACCACTGAGGGGGGACCAACCGGGTCTGGAGACCTC 413
QY 393 CGAGTACCCCGGACCTCCATACCCCATTTCC 425
Db 414 CGAGTACCCCGGACCTCCATACCCCATTTCC 446
```

RESULT 12

LOCUS BQ930450 954 bp mRNA linear EST 20-AUG-2002
DEFINITION AGENCOURT_8946139 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6462494
5', mRNA sequence.

ACCESSION BQ930450
VERSION BQ930450.1 GI:22345481
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 954)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: ILCM2652 row: h column: 15
High quality sequence stop: 576.

FEATURES

SOURCE

Location/Qualifiers

1..954
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6462494"
/clone_lib="NIH_MGC_101"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site_1: EcoRI; Site_2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned

BASE COUNT	205 a	320 c	470 g
ORIGIN			

Query Match	Score	Pred. No.	Indels	Gaps
Best Local Similarity	99.7%	1.3e-83	0	0
Matches	381	Conservative	0	Mismatches 1

OY	44	CAAGAGGCGTCCGAGGAGGAGGAGCCCAATCTCTGTAAGTGGAAAGCTGGGAGGCTCCGAA	103
Db	1	CANAAGGCGTCCGCGAGGAGAGGAGCCCAATCTCTGTAAGTGGAAAGCTGGGAGGCTCCGAA	60
OY	104	GGCCCATCTCAATACCTTCCCTGACTGTGTACAGCCACCTGCTACCCATTTTCTGGAGATGTTTC	163
Db	61	GGCCCATCTCAATACCTTCCCTGACTGTGTGTACAGCCACCTGCTACCCATTTTCTGGAGATGTTTC	120
OY	164	CCCAATATCCACTGGGCTCCCGCGAGCCCCAGTTTGTGGAGGCAAGGGGCGCGCAACAGCA	223
Db	121	CCCAATATCCACTGGGCTCCCGCGAGCCCCAGTTTGTGGAGGCAAGGGGCGCGCAACAGCA	180
OY	224	GAGTCTCTGGAGACCTGACAGGAGACCCGAGTTGGAAACCCGGGGAGCAGAGCCAGGTGCC	283
Db	181	GAGTCTCTCTGGAGACCTGACAGGAGACCCGAGTTGGAAACCCGGGGAGCAGAGCCAGGTGGT	240
OY	284	CCACGGTACCAATGGCATTTCTATGTCAACGGCGGGTCTATGACTATACATGGCAACATCTA	343
Db	241	CCACGGTACCAATGGCATTTCTATGTCAACGGCGGGTCTATGACTATACATGGCAACATCTA	300
OY	344	CATCTACAAATGAGCAAGTACTGGGGGGAGCACCGGGTCTCTGGAGACCTCCAGCTACCCC	403
Db	301	CATCTACAAATGAGCAAGTACTGGGGGGAGCACCGGGTCTCTGGAGACCTCCAGCTACCCC	360
OY	404	CGAAGCTTCATATACCCATTTCCC	425
Db	361	CGAAGCTTCATATACCCATTTCCC	382

RESULT 13					
BG576875					
LOCUS	BG576875	700 bp	mRNA	linear	EST 10-APR-2001
DEFINITION	602595233F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:4707753 5', mRNA sequence.				

ACCESSION	BG576875	GI:13584528
VERSION	BG576875.1	
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	

REFERENCE
1 (bases 1 to 700)
www.ncbi.nlm.nih.gov/

AUTHORS NIH-MDC, <http://mhcni.nlm.nih.gov>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
rs1@scg.rii.nih.gov

Email: cgapuser@comcast.net
 Tissue Procurement: DCTP/NIH
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MCC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://www.llnl.gov/IMG>

http://image.dml.gov
Plate: LLAM10578 row: b column: 10
High quality sequence stop: 673.

```
FEATURES
  source      Location/Vuallierc
              1. .700
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
```

Query Match	87.1%;	Score 3/0;	DB 12;	Length 100;
Best Local Similarity	99.58;	Pred. NO. 4.4e-81;		
Matches 392;	Conservative	0;	Mismatches	0;
			Indels	2;
			Gaps	2

QY	33	GGATGCTGCTCMAAGAGGCGCCGACAGGAGAGGAGACCACCATCTCTGTAACATGAGAC	136
Db	84	GGATGGCTGCTCMAAGAGGCGCTCCGAGGAGAGGAGACCACCATCTCTGTAACATGAGAC	143
QY	93	GAGCCTCGAAGAGGCCATCATCTACTCTGACTTGGTAAAGCAGCTGGTACCATTTCT	152
Db	144	GAGCCTCGAAGAGGCCATCATCTACTCTGACTTGGTAAAGCAGCTGGTACCATTTCT	203
QY	153	GGAGATGTTTCCCGCAGTATCCACTGGGGGTGCCGACGCCAGTTTGGAGGAGGGGTG	212
Db	204	GGAGATGTTTCCCGCAGTATCCACTGGGGTCCCGCAGGCCAGTTTGGAGGAGGGGTG	263
QY	213	CCGCACACAGACAGTCCCTGTGCAGCTGCACAGGAGCCGCAAGTTGGAAACCGGGGAGACAG	272
Db	264	CCGCACACAGACAGTCTGTGCAGCTGTGCACAGGAGCCGCAAGTTGGAAACCGGGGAGACAG	323
QY	273	AGCCAGGTGGCCACGCGTACCAATGGCATCTATGTACACCGGGGTCTATGACTATCACT	332
Db	324	AGCCAGGTGGCCACGCGTACCAATGGCATCTATGTACACCGGGGTCTATGACTATCACT	383
QY	333	GGCAAC-ATTACATCTACAATGAGACAGTACTGGGGGGAGCACCGGGTCTCTGGAGACT	391
Db	384	GGCAACATCTACATCTACAATGAGACAGTACTGGGGGGAGCACCGGGTCTCTGGAGACT	443
QY	392	CCGAGCTACCCCGAACCCTCATACCCCATTTCC	425
Db	444	CCGAGCTA-CCCGAACCCTCATACCCCATTTCC	476

RESULT 14	749 bp	EST 20-OCT-2000
BE789430		
LOCUS	BE789430	
DEFINITION	601482267F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3884920 5', mRNA sequence.	

ACCESSION	BE/89430	GI:10210628
VERSION	BE789430.1	
KEYWORDS	EST,	
SOURCE	human.	

ORGANISM Homo sapiens
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Primates: Catarrhini: Hominiidae: Homo.
1 (bases 1 to 749)

AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
E-mail: nih-gov

Email: c9qapob-1@mail.ru
 Tissue Procurement: DCTD/DTP/Gazdar
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: Invitae Genomics, Inc.
 DNA Sequencing by: Invitae Genomics, Inc.
 Clone distribution: MCC clone distribution information can be
 accessed through the I.M.A.G.E. Consortium/LLNL at:

http://image.lnl.gov
plate: LIAM9659 row: e column: 17
High quality sequence stop: 743.
location/Qualifiers

FEATURES

source 1. .749
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3884920"
/clone_1lb="NIH_MGC_68"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.8 kb. Library constructed by Life
Technologies."
BASE COUNT 167 a 241 c 220 g 121 t
ORIGIN
Query Match 83.5%; Score 355; DB 12; Length 749;
Best Local Similarity 99.7%; Pred. No. 2.2e-77;
Matches 366; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 60 GGAGAGGAGCCAA-TCCTGTAGCTGGAAGCTGGAGAGCTCCGAAGGCCATCCATCTT 118
Db 1 GGAGAGGAGCCAAAGTCTGTAGCTGGAAGCTGGAGAGCTCCGAAGGCCATCCATCTT 60
QY 119 CCCGACTGTGTACAGCCACTGCTACCATTTCTGAGATGTTTCCCATGATCCACTGG 178
Db 61 CCCGACTGTGTACAGCCACTGCTACCATTTCTGAGATGTTTCCCATGATCCACTGG 120
QY 179 GCTCCCGCAGCCCGAGTTTGGAGGAGGAGGCTGCGCAACAGAGAGTCTCTGGAGCT 238
Db 121 GCTCCCGCAGCCCGAGTTTGGAGGAGGAGGCTGCGCAACAGAGAGTCTCTGGAGCT 180
QY 239 GACCAGGAGGCGCAGTTGGAACCCGGGAGACAGAGCCAGGTGGCCACAGGTACCAATGG 298
Db 181 GACCAGGAGGCGCAGTTGGAACCCGGGAGACAGAGCCAGGTGGCCACAGGTACCAATGG 240
QY 299 CATTCATGTCTACCGCGGGGTGTATGACTATCATCTGCGCAACATCTCAATGAGAC 358
Db 241 CATTCATGTCTACCGCGGGGTGTATGACTATCATCTGCGCAACATCTCAATGAGAC 300
QY 359 AGTACTGGGGGAGCACCGGGGTCTGGAGACTCCGAGCTCCCGGCAAGCTCCATGCC 418
Db 301 AGTACTGGGGGAGCACCGGGGTCTGGAGACTCCGAGCTCCCGGCAAGCTCCATGCC 360
QY 419 CATTCGCC 425
Db 361 CATTCGCC 367
RESULT 15
BE207591 579 bp mRNA linear EST 27-JUN-2000
LOCUS BE207591
DEFINITION bb66f10.y1 NIH_MGC.9 Homo sapiens cDNA clone IMAGE:3030475 5'
similar to gb:U04270 TUMOR NECROSIS FACTOR RECEPTOR 2 RELATED
PROTEIN PRECURSOR (HUMAN); gb:U38423 Mus musculus lymphotoxin-beta
receptor gene, complete (MOUSE);, mRNA sequence.
ACCESSION BE207591
VERSION BE207591.1 GI:8750989
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 579)
NIH-MGC http://mgc.ncl.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:

image.lnl.gov/image/htnl/lresources.shtml
Seq primer: -40RP from Gibco
High quality sequence stop: 427.
Location/Qualifiers
source 1. .579
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3030475"
/clone_1lb="NIH_MGC_9"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dt priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGCAAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 100 a 198 c 152 g 127 t 2 others
ORIGIN
Query Match 81.6%; Score 346.6; DB 10; Length 579;
Best Local Similarity 92.1%; Pred. No. 2.4e-75;
Matches 386; Conservative 0; Mismatches 31; Indels 2; Gaps 2;
QY 8 AGAGCCACTGCCCCAGAGATGTCAGATGCTCTCAAGAGGCTCCGAGGAGAGGG 67
Db 20 AGAGCCACTGCTCTCTGCGGAATGAGATGCTCTCTCAAGA-GGCTCCGAGGAGAGGG 78
QY 68 ACCCAATCTGTAGCTGGAAGCTGGAGCCCTCCGAAGGCCCATCACTTCCCTGACTT 127
Db 79 ACCCAATCTGTGTCTGGAAGCTGGAGCCCTCCGAAGGCCCATCACTTCTCTGACTT 138
QY 128 GGTACAGCAGCTGCTACCAT-TTCTGAGAGATGTTTCCCGATATCACTGGGCTCCCG 186
Db 139 GGTACAGCAGCTGTTGCCCGGTGTGGAGATGTTTCCCGATATCACTGGGCTCCCG 198
QY 187 CAGCCCAAGTTTGGAGGAGGAGGCTCCGAACAGAGAGTCTCTGAGCTGACAGGG 246
Db 199 CAGCCCAAGTTTGGAGGAGGAGGCTCCGAACAGAGAGTCTCTGAGCTGACAGGG 258
QY 247 AGCCGAGTTGGAACCCGGGAGAGAGCCAGTGGCCACAGGTACCAATGGCAATTCATG 306
Db 259 AGCCGAGTTGGAACCCGGGAGAGAGCCAGTGGCCACAGGTACCAATGGCAATTCATG 318
QY 307 TCACCGGGGGGTCTATGACTATCACTGAGCAACATCTACATCAATGAGACAGTACTGG 366
Db 319 TCACCGGGGGGTCTATGACTATCACTGAGCAACATCTCTACATGAGACATTTCTGG 378
QY 367 GGGAGCAGCCGGGTCTGAGAGACCTCCAGTATACCCGCAATCCCATTTCC 425
Db 379 GGGAGCAGCCGGGTCTGAGAGACCTCCAGTATACCCGCAATCCCATTTCC 437
Search completed: April 15, 2003, 16:26:35
Job time : 495.898 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:11:20 ; Search time 13.6946 Seconds
(without alignments)
9517.464 Million cell updates/sec

Title: US-09-917-372-2_COPY_853_1277
Perfect score: 425
Sequence: 1 atccatgagccactgcc.....aacctcatcaccatccc 425

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/PTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/Dackfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34.2	8.0	268	1	US-08-039-137-16 Sequence 16, Appl
2	33	7.8	1057	3	US-09-188-930-18 Sequence 18, Appl
3	33	7.8	1722	4	US-09-142-623-14 Sequence 14, Appl
4	33	7.8	2004	3	US-09-188-930-230 Sequence 230, App
5	32.8	7.7	376	4	US-09-149-476-201 Sequence 201, App
6	32.2	7.6	2861	4	US-08-482-073-10 Sequence 10, Appl
7	32.2	7.6	3647	1	US-07-914-281-7 Sequence 7, Appl
8	32.2	7.6	3647	1	US-08-393-246-7 Sequence 7, Appl
9	32.2	7.6	3647	1	US-08-525-058A-7 Sequence 7, Appl
10	32.2	7.6	3647	2	US-08-696-731-7 Sequence 7, Appl
11	32.2	7.6	3647	4	US-09-042-531-7 Sequence 7, Appl
12	32.2	7.6	3647	5	PCT-US91-00899-4 Sequence 4, Appl
13	32.2	7.6	4403765	4	US-09-103-840A-2 Sequence 2, Appl
14	31.8	7.5	1977	4	US-09-548-372D-5 Sequence 5, Appl
15	31.8	7.5	1977	4	US-09-548-367D-5 Sequence 5, Appl
16	31.8	7.5	2070	4	US-09-548-372D-3 Sequence 3, Appl
17	31.8	7.5	2070	4	US-09-548-367D-3 Sequence 3, Appl
18	31.8	7.5	36741	4	US-09-301-665-3 Sequence 3, Appl
19	31.6	7.4	550	4	US-09-149-476-19 Sequence 19, Appl
20	31	7.3	269	1	US-08-050-073-1 Sequence 1, Appl
21	31	7.3	269	1	US-08-050-073-2 Sequence 2, Appl
22	31	7.3	269	1	US-08-050-073-11 Sequence 11, Appl
23	31	7.3	269	1	US-08-050-073-14 Sequence 14, Appl
24	31	7.3	269	1	US-08-050-073-38 Sequence 38, Appl
25	31	7.3	1244	3	US-08-463-903-23 Sequence 23, Appl
26	31	7.3	1244	4	US-07-935-695-23 Sequence 23, Appl
27	31	7.3	4403765	4	US-09-103-840A-2 Sequence 2, Appl

28	30.4	7.2	1930	4	US-09-724-864-9 Sequence 9, Appl
29	30.4	7.2	1953	3	US-08-826-246-1 Sequence 1, Appl
30	30.4	7.2	1953	3	US-08-944-495-1 Sequence 1, Appl
31	30.4	7.2	1953	3	US-09-126-640-1 Sequence 1, Appl
32	30.4	7.2	1953	4	US-08-925-588-1 Sequence 1, Appl
33	30.4	7.2	1953	4	US-09-288-292A-1 Sequence 1, Appl
34	30.4	7.2	4086	3	US-08-313-181-1 Sequence 3, Appl
35	30.4	7.2	43950	4	US-09-735-934A-3 Sequence 3, Appl
36	30.2	7.1	2070	1	US-08-394-326-1 Sequence 1, Appl
37	30.2	7.1	2070	3	US-09-082-306-1 Sequence 1, Appl
38	30	7.1	1710	1	US-07-903-103-3 Sequence 3, Appl
39	30	7.1	1710	1	US-08-044-619A-3 Sequence 3, Appl
40	30	7.1	1710	1	US-08-283-911-3 Sequence 3, Appl
41	30	7.1	1710	1	US-08-245-500A-4 Sequence 4, Appl
42	30	7.1	1710	1	US-08-390-546-4 Sequence 4, Appl
43	30	7.1	1710	1	US-08-390-479A-4 Sequence 4, Appl
44	30	7.1	1710	1	US-08-357-393-4 Sequence 4, Appl
45	30	7.1	1710	1	US-08-390-516C-4 Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-039-137-16
Sequence 16, Application US/08039137
Patent No. 5759771
GENERAL INFORMATION:
APPLICANT: Tilius J.G., Marcel
TITLE OF INVENTION: Method of Determining a Genotype by
TITLE OF INVENTION: Comparing the Nucleotide Sequence of Members of a Gene
Patent No. 5759771
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSER: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/039,137
FILING DATE: 14-APR-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 0550-0024.10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 268 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: DRI CDNA, Fig. 1
FEATURE:
NAME/KEY: CDS
LOCATION: 1..267
US-08-039-137-16

```
Query Match      8.0%; Score 34.2; DB 1; Length 268;
Best Local Similarity 54.3%; Pred. No. 0.5;
Matches 69; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 207 GGGGCGCCGACACAGAGATCTCTGTGACCTGACAGGAGCCGACGTTGGAAACCCGGG 266
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 130 GCGGTGACGAGACTGGGGCGGGCTGATGCCGAGTACTGGAACGACAGAACGACCTCTG 189

QY 267 GAGCAGAGCCAGGTGGCCGACGAGTACCAATGCACTATGTCACCGGGGCTATGACT 326
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 190 GAGCAGAGCGCGCGCGGTGACACCTACTGACAGACACAACTACGCGGTGGTAGAGC 249

QY 327 ATCACTG 333
      |||||
Db 250 TTCACAG 256

RESULT 2
US-09-188-930-18
: Sequence 18, Application US/09188930A
: Patent No. 6150502
: GENERAL INFORMATION:
: APPLICANT: Watson, James D.
: APPLICANT: Strachan, Lorna
: APPLICANT: Sleeman, Matthew
: APPLICANT: Murison, Rene
: TITLE OF INVENTION: Compositions Isolated From Skin Cells
: FILE REFERENCE: 11000.1011c1
: CURRENT APPLICATION NUMBER: US/09/188,930A
: CURRENT FILING DATE: 1998-11-09
: NUMBER OF SEQ ID NOS: 348
: SOFTWARE: PastSeq for Windows Version 3.0
: SEQ ID NO 18
: LENGTH: 1057
: TYPE: DNA
: ORGANISM: Rat
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (45)...(45)
: NAME/KEY: unsure
: LOCATION: (53)...(53)
: NAME/KEY: unsure
: LOCATION: (116)...(116)
: NAME/KEY: unsure
: LOCATION: (118)...(118)
: US-09-188-930-18

Query Match      7.8%; Score 33; DB 3; Length 1057;
Best Local Similarity 51.0%; Pred. No. 1.7;
Matches 78; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 53 TCCGAGGGAAGGAGCCCAATCTGTAGCTGGAAGCTGGGAGCCCTCCGAAGCCCATCC 112
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 353 TCCGAGCGCTGTGGATCCGATCGCAGCGGAGGAGTGCTCCATTAACGAGCTTCC 412

QY 113 ATACTCCCTGACTTGTGACAGCCACTGTACCATTTCTGGAATGTTTCCCACTATC 172
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 413 CTAAAGCCCTGGACAGTGTGATCTCCAAAGTACTCTCTCGGGAGCTACACCTGTCAATCA 472

QY 173 CACTGGGCTCCCGAGCCGAGTGGAGGC 205
      ||||| ||||| ||||| |||||
Db 473 CGCAAGGCTTTGGAGGACCGGATATCTGGGGC 505

RESULT 3
US-09-142-623-14/c
: Sequence 14, Application US/09142623
: Patent No. 6337201
: GENERAL INFORMATION:
: APPLICANT: Koji YANAI et al.
: TITLE OF INVENTION: -FRUCTOFURANOSIDASE AND ITS GENE, METHOD OF
```

```

: TITLE OF INVENTION: ISOLATING -FRUCTOFURANOSIDASE GENE, SYSTEM FOR PRODUCING
: TITLE OF INVENTION: -FRUCTOFURANOSIDASE, AND -FRUCTOFURANOSIDASE VARIANT
: NUMBER OF SEQUENCES: 35
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
: STREET: 2033 K Street, N.W., Suite 800
: CITY: Washington
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20006
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: MS-DOS
: SOFTWARE: Wordperfect 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/142,623
: FILING DATE: September 10, 1998
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Lee Cheng
: REGISTRATION NUMBER: 40,949
: REFERENCE/DOCKET NUMBER: 98-0989*/LC(WMC)/144
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-721-8200
: TELEFAX: 202-721-8250
: INFORMATION FOR SEQ ID NO: 14:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1722 base pairs
: TYPE: Nucleic acid
: STRANDEDNESS: Double stranded
: TOPOLOGY: Linear
: MOLECULE TYPE: Genomic DNA
: ORGANISM: Microorganism: Scopulariopsis brevicaulis IF04843
: FEATURE:
: NAME/KEY: mat peptide
: LOCATION: 1...1722
: IDENTIFICATION METHOD: E
: US-09-142-623-14

Query Match      7.8%; Score 33; DB 4; Length 1722;
Best Local Similarity 53.5%; Pred. No. 2;
Matches 69; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 296 TGGCATTCATGTCAACCGCGGGTCTATGACTATGCACTGCAACATCTACATCTCAATGG 355
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 838 TGGCAAAAGATTTCACCGTCGGGGTGTGAGCCATATGCTCAAGAACGACAAATGTTGATGA 779

QY 356 ACCAGTACTGGGGGAGACACCGGGGCTCTGGAGACTCCCACTACCCCGGAACCTCCATA 415
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 778 CCTGAAATTTTACCCCAACCGGAGCCAGTCACCGCTGCCCAAGTGAGTCCCTT 719

QY 416 CCCCATTC 424
      |||||
Db 718 CCTGCTTC 710

RESULT 4
US-09-188-930-230
: Sequence 230, Application US/09188930A
: Patent No. 6150502
: GENERAL INFORMATION:
: APPLICANT: Watson, James D.
: APPLICANT: Strachan, Lorna
: APPLICANT: Sleeman, Matthew
: APPLICANT: Murison, Rene
: TITLE OF INVENTION: Compositions Isolated From Skin Cells
: FILE REFERENCE: 11000.1011c1
```

```

; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 230
; LENGTH: 2004
; TYPE: DNA
; ORGANISM: Rat
US-09-188-930-230

Query Match      7.8%; Score 33; DB 3; Length 2004;
Best Local Similarity 51.0%; Pred. No. 2.1;
Matches 78; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY  53  TCCGCAGAGAGAGGAGCCATCTGTACTGTGAGAGCCCTCCGAGGCCATCC 112
      |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB  287  TCCGCACGGGTGGATTCGATCTGCAGCGAGAGAGTGTCCATTACAGGCTTTCC 346
      ||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
QY  113  ATACTTCCCTGACTGTGTACAGCCACTGCTACCCATTCTGTGAGATGTTCCCACTATC 172
      ||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
DB  347  CTAAAGCCCTGGAGACGTGTGATCTCCAGTACTCTGTGGGAGACTACACCTGTATTC 406
      ||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
QY  173  CACTGGGCTCCCGCAGCCCACTTTTGAGGC 205
      ||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
DB  407  CGCAAGCTTTTGAGAGCCCGATCTGAGGC 439
      ||  |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||

RESULT 5
US-09-149-476-201/c
; Sequence 201, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
```

```

; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
```

```

? EARLIER APPLICATION NUMBER: 60/057,650
? EARLIER FILING DATE: 1997-09-05
? EARLIER APPLICATION NUMBER: 60/056,884
? EARLIER FILING DATE: 1997-08-22
? EARLIER APPLICATION NUMBER: 60/057,669
? EARLIER FILING DATE: 1997-09-05
? EARLIER APPLICATION NUMBER: 60/049,610
? EARLIER FILING DATE: 1997-06-13
? EARLIER APPLICATION NUMBER: 60/061,060
? EARLIER FILING DATE: 1997-10-02
?
Query_Match      7.7%;    Score 32.8; DB 4; Length 376;
Best Local Similarity 47.5%; Pred.No.1.4; Indels 0; Gaps 0;
Matches 94; Conservative 1; Mismatches 103;

```

RESULT 6

US-08-482,073 (7/96)
Sequence 10, Application US/08482073
Patent No. 6307025
GENERAL INFORMATION:
APPLICANT: Hession, Catherine A.
APPLICANT: Lobb, Roy R.
APPLICANT: Goelz, Susan E.
APPLICANT: Osborn, Laurelee
APPLICANT: Benjamin, Christopher D.
APPLICANT: Rosa, Margaret D.
TITLE OF INVENTION: ENDOTHELIAL CELL-LEUKOCYTE ADHESION
TITLE OF INVENTION: MOLECULES (ELAMS) AND MOLECULES INVOLVED IN LEUKOCYTE
TITLE OF INVENTION: ADHESION (MILAS)
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Neave
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,073
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/486,336
FILING DATE:
APPLICATION NUMBER: US 07/608298
FILING DATE: 31-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US 90/02357
FILING DATE: 27-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/452675

FILING DATE: 18-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/359516
FILING DATE: 01-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/354151
FILING DATE: 28-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Haley Jr., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B124C1P4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 2861 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-482-073-10

Query Match 7.6%; Score 32.2; DB 4; Length 2861;
Best Local Similarity 57.4%; Pred. No. 4;
Matches 58; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 179 GCTCCCCGAGCCCGAGTTTGGAGGAGCGGTGCCGCAACAGCAGATCTCTGACCT 238
DB 447 GCTCGCGCGTGGCGCCGGAGTGTGAAGGGGCGGTCCCAAGATGCAAGGCGCGGCGCT 388
QY 239 GACCAAGGAGCGCGAGTTGGAAACCCGGGAGGAGCAGAGCCAGG 279
DB 387 GCCCTGCTCTCCCAAGTGCAGGCGGGGCGGAGCGCCAGG 347

RESULT 7

US-07-914-281-7/C
Sequence 7, Application US/07914281
Patent No. 5324663
GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCT
NUMBER OF SEQUENCES: 14
CURRENT APPLICATION DATA:
CORRESPONDENCE ADDRESS:
ADDRESSER: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESSER: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/914,281
FILING DATE: 19920720
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:
LENGTH: 3647 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: NO
US-07-914-281-7

Query Match 7.6%; Score 32.2; DB 1; Length 3647;
Best Local Similarity 57.4%; Pred. No. 4.3;
Matches 58; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 179 GCTCCCCGAGCCCGAGTTTGGAGGAGCGGTGCCGCAACAGCAGATCTCTGACCT 238
DB 1840 GCTCGCGCGTGGCGCCGGAGTGTGAAGGGGCGGTCCCAAGATGCAAGGCGCGGCGCT 1781
QY 239 GACCAAGGAGCGCGAGTTGGAAACCCGGGAGGAGCAGAGCCAGG 279
DB 1780 GCCCTGCTCTCCCAAGTGCAGGCGGGGCGGAGCGCCAGG 1740

RESULT 8

US-08-393-246-7/C
Sequence 7, Application US/08393246
Patent No. 5595900
GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCT
NUMBER OF SEQUENCES: 14
CURRENT APPLICATION DATA:
CORRESPONDENCE ADDRESS:
ADDRESSER: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESSER: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/393,246
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/220,433
FILING DATE: 30-MAR-1994
APPLICATION NUMBER: US 07/914,281
FILING DATE: 20-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3647 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: NO
US-08-393-246-7

Query Match 7.6%; Score 32.2; DB 1; Length 3647;
Best Local Similarity 57.4%; Pred. No. 4.3;
Matches 58; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

OY 179 GCTCCCGCAGCCCGCAGTTTGGAGCGAGGTCCTCTGAGCT 238
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 1840 GCTGCGCGCTGCGCGGGAATGGAAGGGGCTCCCAATGCAAGGCCGCGGCCCT 1781

OY 239 GACCAAGGAGCGCGAGTTGGAACCCGGGAGCAGAGCCAGG 279
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 1780 GCCCTGCTCTCCCAAGTGGCGGGGGGAGCGCCAGG 1740

RESULT 9
US-08-525-058A-7/c
Sequence 7, Application US/08525058A
Patent No. 5770420
GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: OHLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/525,058A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3647 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: NO

US-08-525-058A-7

Query Match 7.6%; Score 32.2; DB 1; Length 3647;
Best Local Similarity 57.4%; Pred. No. 4.3;
Matches 58; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

OY 179 GCTCCCGCAGCCCGCAGTTTGGAGCGAGGTCCTCTGAGCT 238
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 1840 GCTGCGCGCTGCGCGGGAATGGAAGGGGCTCCCAATGCAAGGCCGCGGCCCT 1781

OY 239 GACCAAGGAGCGCGAGTTGGAACCCGGGAGCAGAGCCAGG 279
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 1780 GCCCTGCTCTCCCAAGTGGCGGGGGGAGCGCCAGG 1740

RESULT 10
US-08-696-731-7/c

Sequence 7, Application US/08696731
Patent No. 595347
GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 14-AUG-1996
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: 08/393,246
APPLICATION NUMBER: 08/393,246
FILING DATE:
APPLICATION NUMBER: US 08/220,433
FILING DATE: 30-MAR-1994
APPLICATION NUMBER: US 07/914,281
FILING DATE: 20-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3647 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: NO

US-08-696-731-7

Query Match 7.6%; Score 32.2; DB 2; Length 3647;
Best Local Similarity 57.4%; Pred. No. 4.3;
Matches 58; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

OY 179 GCTCCCGCAGCCCGCAGTTTGGAGCGAGGTCCTCTGAGCT 238
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 1840 GCTGCGCGCTGCGCGGGAATGGAAGGGGCTCCCAATGCAAGGCCGCGGCCCT 1781

OY 239 GACCAAGGAGCGCGAGTTGGAACCCGGGAGCAGAGCCAGG 279
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 1780 GCCCTGCTCTCCCAAGTGGCGGGGGGAGCGCCAGG 1740

RESULT 11
US-09-042-531-7/c
Sequence 7, Application US/09042531
Patent No. 6268193
GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION

```

Query Match          7.6%; Score 32.2; DB 4; Length 3647;
Best Local Similarity 57.4%; Pred. No. 4.3;
Matches 58; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

0Y 179 GCTCCCGCAGCCCCAGTTTGGAGCGAGGGGTCGCCCAACAGCAGAGTCTCTGAGACT 238
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1840 GCTGCGCGCTGCCCGGGAATGGAAGGGGGGCGTCCCAAGATGCAAAAGCGCCGCGCCT 1781

0Y 239 GACCAGGAGCGCGCAGTTGCAACCCGGGGAGACAGAGCCAGG 279
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1780 GCCTGCTCTCTCCCAAGTGGCGGGGGGAGCGCCGACAG 1740

RESULT 12
Sequence 4, Application PC/TUS9100899
GENERAL INFORMATION:
APPLICANT: Lowe, John B.
TITLE OF INVENTION: Method and Products for the Synthesis of
TITLE OF INVENTION: Oligosaccharide Structures on Glycoproteins, Glycolipids
or as Free Molecules, and for the Isolation of Cloned
TITLE OF INVENTION: Genetic Sequences That Determine These Structures
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCQUELLAND, MATER & NEUSTADT,
ADDRESS: P. C.
STREET: 1755 Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia

```

Query Match	7.6%	Score 32.2	DB 5	Length 3647
Best Local Similarity	57.4%	Pred. No. 4,3		
Matches 58	Conservative 0	Mismatches 43	Indels 0	Gaps 0
QY 179	GCTCCCCGACGCCAGTTTGGAGGCGACAGGGGTGCCCCACAGCAGAGTCTCTGGACCT			238
Db 1840	GCTCCCTCTGCTGCCCCGGAAGTGGAGGGGGCGCTCCAGATGCAAGAGCCCGGGGCTT			1781
QY 239	GACCAAGGAGGCCGCGCAGTTGGAAACCCGGGGAGAGCAGAGCCAGG			279
Db 1780	GCCCTGCTCTCTCCCAAGTGGCGGGGGGGGCGAGCCGACAGG			1740

```

RESULT 13
US-09-103-840A-2/C
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: TUBERCULOSIS
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match          7.6%; Score 32.2; DB 4; Length 4403765;
Best Local Similarity 48.9%; Pred. No. 31;
Matches 114; Conservative 0; Mismatches 118; Indels 1; Gaps 1.;

OY      129 GTACAGCCACTGCTACCACTTTCTGGAGATGTTCCTCCACAGTATCCACTGCGGCTCCCCGCCA 188
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
GC3934365 GCANAAGCAGAGTACCATTGCGCGGAGCGCTCGGGGCGGTGCGGCGCCGCGCACTGCGC 3934306

```

```

RESULT 15
US-09-548-367D-5
; Sequence 5, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
; APPLICANT: GUREY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/6280H
; CURRENT APPLICATION NUMBER: US/09/548,367D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23

```

Search completed: April 15, 2003, 16:46:55
Job time : 1040.69 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 15:24:11 ; Search time 29.6131 Seconds
(without alignments)
12588.918 Million cell updates/sec

Title: US-09-917-372-2_COPY_853_1277

Perfect score: 425

Sequence: 1 atccattagagccacgtcc.....aacctcataccattccc 425

Scoring table: IDENTITY NUC

Gapop:10.0 , Gapext 1.0

Searched: 593429 seqs, 438583890 residues

Total number of hits satisfying chosen parameters: 1186858

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCR_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCR_NEW_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	425	100.0	425	10	US-09-907-372-9
2	425	100.0	1982	10	US-09-907-372-2
3	423.4	99.6	574	10	US-09-907-372-8
4	403	94.8	651	10	US-09-907-372-7
5	393	92.5	2136	9	US-09-954-531-1348
6	393	92.5	2136	10	US-09-962-436-262
7	393	92.5	2136	10	US-09-880-107-2135
8	253.8	59.1	281	10	US-09-867-701-5601
9	251.2	59.1	289	10	US-09-867-701-5666
10	171.8	40.4	371	10	US-09-907-372-16
11	135.8	32.0	206	10	US-09-907-372-13
12	98.4	23.2	219	10	US-09-907-372-10
13	71	16.7	279	10	US-09-907-372-11
14	47.8	11.2	862	10	US-09-907-372-12
15	35.4	8.3	809	10	US-09-844-864-22
16	34	8.0	6788	9	US-09-870-759-101
17	33.6	7.9	33675	10	US-09-921-992-2
18	33	7.8	1057	9	US-10-152-661-18
19	33	7.8	1057	9	US-09-866-050A-18

C	20	33	7.8	1722	9	US-09-990-385-14	Sequence 14, Appl
	21	33	7.8	1890	9	US-10-152-661-447	Sequence 447, App
	22	33	7.8	1890	9	US-09-866-050A-447	Sequence 447, App
	23	33	7.8	2004	9	US-10-152-661-230	Sequence 230, App
	24	33	7.8	2004	9	US-09-866-050A-230	Sequence 230, App
C	25	32.8	7.7	376	9	US-09-809-391-201	Sequence 201, App
	26	32.2	7.6	785	9	US-09-902-941-1868	Sequence 1868, Ap
	27	32.2	7.6	785	9	US-09-849-626-1868	Sequence 1868, Ap
	28	32.2	7.6	785	9	US-10-017-754-1868	Sequence 1868, Ap
C	29	32.2	7.6	3647	10	US-09-863-475A-7	Sequence 7, Appl
	30	31.8	7.5	671	9	US-10-184-644-346	Sequence 340, App
	31	31.8	7.5	1911	9	US-09-470-954A-44	Sequence 44, Appl
	32	31.8	7.5	1977	10	US-09-794-927-5	Sequence 5, Appl
	33	31.8	7.5	1977	10	US-09-794-927-5	Sequence 5, Appl
	34	31.8	7.5	1977	10	US-09-794-927-5	Sequence 5, Appl
	35	31.8	7.5	1977	10	US-09-794-927-5	Sequence 5, Appl
	36	31.8	7.5	1977	10	US-09-794-927-5	Sequence 5, Appl
	37	31.8	7.5	1977	10	US-09-681-442-5	Sequence 5, Appl
	38	31.8	7.5	2070	10	US-09-794-927-3	Sequence 3, Appl
	39	31.8	7.5	2070	10	US-09-795-847-3	Sequence 3, Appl
	40	31.8	7.5	2070	10	US-09-794-743-3	Sequence 3, Appl
	41	31.8	7.5	2070	10	US-09-794-748-3	Sequence 3, Appl
	42	31.8	7.5	2070	10	US-09-794-925-3	Sequence 3, Appl
	43	31.8	7.5	2070	10	US-09-681-442-3	Sequence 3, Appl
	44	31.8	7.5	3252	9	US-09-795-903A-1	Sequence 1, Appl
	45	31.8	7.5	3252	10	US-09-796-264-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-907-372-9/C
Sequence 9, Application US/09907372
Patent No. US20020068242A1
GENERAL INFORMATION:
APPLICANT: Lal, Preethi G.
TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
FILE REFERENCE: PC-0050 US
CURRENT APPLICATION NUMBER: US/09/907 372
CURRENT FILING DATE: 2001-07-27
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PERL Program
SEQ ID NO 9
LENGTH: 425
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID NO. US20020068242A1 7716340H1
US-09-907-372-9

Query Match	100.0%;	Score 425;	DB 10;	Length 425;
Best Local Similarity	100.0%;	Pred. No. 6.5e-124;		
Matches	425;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
OY	1	ATCCATTAGAGCCACTGCCCCGAGAGATGTCAGATGCTGTCACAGAGCGTCGCGAGG	60	
DB	425	ATCCATTAGAGCCACTGCCCCGAGAGATGTCAGATGCTGTCACAGAGCGTCGCGAGG	366	
OY	61	GAGAGGAGCCCAATCCTGAGCTGGAAGCTGGAGAGCTCCGAAGGCCCATCATCTTCC	120	
DB	365	GAGAGGAGCCCAATCCTGAGCTGGAAGCTGGAGAGCTCCGAAGGCCCATCATCTTCC	306	
OY	121	CTGACTTGATAGAGCCACTGATCCATTTCTGAGATGTTTCCCATGATATCCACTGGGC	180	
DB	305	CTGACTTGATAGAGCCACTGATCCATTTCTGAGATGTTTCCCATGATATCCACTGGGC	246	
OY	181	TCGCCGACAGCCCACTTTTGGAGGAGGGGTCGCCACAGACAGAGATCTTGACCTGA	240	
DB	245	TCGCCGACAGCCCACTTTTGGAGGAGGGGTCGCCACAGAGATCTTGACCTGA	186	

```

OY 241 CCAGGAGCCGCGAGTTGGAACCCGCGGAGACGACGAGTGGGCCGAGTACCAATGGA 300
    |||||
Db 185 CCAGGAGCCGCGAGTTGGAACCCGCGGAGACGACGAGTGGGCCGAGTACCAATGGA 126
OY 301 TTCATGTCACCGCGGGGTCTATGACTATCTAGTGGCAACATCTACATCTAACATGAGCAG 360
    |||||
Db 125 TTCATGTCACCGCGGGGTCTATGACTATCTAGTGGCAACATCTACATCTAACATGAGCAG 66
OY 361 TACTGGGGGAGCACCAGGGTCTCGTGAGACCTCCGAGCTCCCGGACCTCCATACCCCA 420
    |||||
Db 65 TACTGGGGGAGCACCAGGGTCTCGTGAGACCTCCGAGCTCCCGGACCTCCATACCCCA 6
OY 421 TTCCC 425
    |||||
Db 5 TTCCC 1

```

RESULT 2

```

US-09-907-372-2
; Sequence 2, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 1982
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 7497867CBI
US-09-907-372-2

```

Query Match

```

Best Local Similarity 100.0%; Score 425; DB 10; Length 1982;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 ATCCATTAGAGCCACTGCCCCCAAGATGTGAGATGCTGCTCTCAAGAGCGTCCGAGG 60
    |||||
Db 853 ATCCATTAGAGCCACTGCCCCCAAGATGTGAGATGCTGCTCTCAAGAGCGTCCGAGG 912
OY 61 GAGAGGAGCCCAATCTGTAGCTGGAAGCTGGAGCGCTCGAAGGCCCATCCATCTCC 120
    |||||
Db 913 GAGAGGAGCCCAATCTGTAGCTGGAAGCTGGAGCGCTCGAAGGCCCATCCATCTCC 972
OY 121 CTGACTTGATACAGCCACTGCTACCAATTTCTGAGATGTTTCCCAATATCCACTGGG 180
    |||||
Db 973 CTGACTTGATACAGCCACTGCTACCAATTTCTGAGATGTTTCCCAATATCCACTGGG 1032
OY 181 TCCCGGAGGCCAGTTTGGAGGCGGGTCCGCAACAGCAGAGTCTCTGGAGCTGA 240
    |||||
Db 1033 TCCCGGAGGCCAGTTTGGAGGCGGGTCCGCAACAGCAGAGTCTCTGGAGCTGA 1092
OY 241 CCAGGAGCCGCGAGTTGGAACCCGCGGAGACGAGCCAGTGGGCCAGGTAACCAATGGCA 300
    |||||
Db 1093 CCAGGAGCCGCGAGTTGGAACCCGCGGAGACGAGCCAGTGGGCCAGGTAACCAATGGCA 1152
OY 301 TTCATGTCACCGCGGGGTCTATGACTATCTAGTGGCAACATCTACATCTAACATGAGCAG 360
    |||||
Db 1153 TTCATGTCACCGCGGGGTCTATGACTATCTAGTGGCAACATCTACATCTAACATGAGCAG 1212
OY 361 TACTGGGGGAGCACCAGGGTCTCTGAGACCTCCAGCTACCCCGGACCTCCATACCCCA 420
    |||||
Db 1213 TACTGGGGGAGCACCAGGGTCTCTGAGACCTCCAGCTACCCCGGACCTCCATACCCCA 1272
OY 421 TTCCC 425
    |||||

```

Db 1273 TTCCC 1277

RESULT 3

```

US-09-907-372-8
; Sequence 8, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 8
; LENGTH: 574
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 8234468H1
US-09-907-372-8

```

Query Match

```

Best Local Similarity 99.8%; Score 423.4; DB 10; Length 574;
Matches 424; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 ATCCATTAGAGCCACTGCCCCCAAGATGTGAGATGCTGCTCTCAAGAGCGTCCGAGG 60
    |||||
Db 50 ATCCATTAGAGCCACTGCCCCCAAGATGTGAGATGCTGCTCTCAAGAGCGTCCGAGG 109
OY 61 GAGAGGAGCCCAATCTGTAGCTGGAAGCTGGAGCGCTCGAAGGCCCATCCATCTCC 120
    |||||
Db 110 GAGAGGAGCCCAATCTGTAGCTGGAAGCTGGAGCGCTCGAAGGCCCATCCATCTCC 169
OY 121 CTGACTTGATACAGCCACTGCTACCAATTTCTGAGATGTTTCCCAATATCCACTGGG 180
    |||||
Db 170 CTGACTTGATACAGCCACTGCTACCAATTTCTGAGATGTTTCCCAATATCCACTGGG 229
OY 181 TCCCGGAGGCCAGTTTGGAGGCGGGTCCGCAACAGCAGAGTCTCTGGAGCTGA 240
    |||||
Db 230 TCCCGGAGGCCAGTTTGGAGGCGGGTCCGCAACAGCAGAGTCTCTGGAGCTGA 289
OY 241 CCAGGAGCCGCGAGTTGGAACCCGCGGAGACGAGCCAGTGGGCCAGGTAACCAATGGCA 300
    |||||
Db 290 CCAGGAGCCGCGAGTTGGAACCCGCGGAGACGAGCCAGTGGGCCAGGTAACCAATGGCA 349
OY 301 TTCATGTCACCGCGGGGTCTATGACTATCTAGTGGCAACATCTACATCTAACATGAGCAG 360
    |||||
Db 350 TTCATGTCACCGCGGGGTCTATGACTATCTAGTGGCAACATCTACATCTAACATGAGCAG 409
OY 361 TACTGGGGGAGCACCAGGGTCTCTGAGACCTCCAGCTACCCCGGACCTCCATACCCCA 420
    |||||
Db 410 TACTGGGGGAGCACCAGGGTCTCTGAGACCTCCAGCTACCCCGGACCTCCATACCCCA 469
OY 421 TTCCC 425
    |||||
Db 470 TTCCC 474

```

RESULT 4

```

US-09-907-372-7/c
; Sequence 7, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27

```

```

; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 651
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 7716364H1
US-09-907-372-7

```

```

Query Match          94.8%; Score 403; DB 10; Length 651;
Best Local Similarity 99.5%; Pred. No. 5,7e-117;
Matches 425; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

```

```

QY 1 ATCCATTNAGCCACTGCCCCCAGAGATGTCAGAGATCCGTCTCAAGAGGCGTCCGACG 60
DB 427 ATCCATTNAGCCACTGCCCCCAGAGATGTCAGAGATCCGTCTCAAGAGGCGTCCGACG 368
QY 61 GAGAGGAGCCCAATCCTGTAGCTGGAAGCTGGAGGCTCCGAAGGCCCATCCATACCTTC 120
DB 367 GAGAGGAGCCCAATCCTGTAGCTGGAAGCTGGAGGCTCCGAAGGCCCATCCATACCTTC 308
QY 121 CTGACTTGTAGACCACTGCTACCACTTCTGAGATGTTTCCCAAGTATCCACTGAGC 180
DB 307 CTGACTTGTAGACCACTGCTACCACTTCTGAGATGTTTCCCAAGTATCCACTGAGC 248
QY 181 TCCCGGACCCCACTTTTGGAGGAGGAGGCGCCGACAGACAGAGTCTCTGGACCTGA 240
DB 247 TCCCGGACCCCACTTTTGGAGGAGGAGGCGCCGACAGACAGAGTCTCTGGACCTGA 188
QY 241 CCAGGAGGCGGAGTGTGAACCCGGGAGCAGAGCCAGGTGGGCCCAAGTACCAATGGA 300
DB 187 CCAGGAGGCGGAGTGTGAACCCGGGAGCAGAGCCAGGTGGGCCCAAGTACCAATGGA 128
QY 301 TTCATGTCAACCGGCGGTCTATGATCATCTGCAATCTACATCTCAATGAGACGAG 360
DB 127 TTCATGTCAACCGGCGGTCTATGATCATCTGCAATCTACATCTCAATGAGACGAG 68
QY 361 TACTGGGGGAGCACCAGGCTCTCGAGACCTCC-AGTACCCCCGAGACCTCC-ATACCC 418
DB 67 TACTGGGGGAGCACCAGGCTCTCGAGACCTCC-AGTACCCCCGAGACCTCC-ATACCC 8
QY 419 CATTCOC 425
DB 7 CATTCOC 1

```

```

RESULT 5
US-09-954-531-1348
; Sequence 1348, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cano
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1348
; LENGTH: 2136

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-531-1348

```

```

Query Match          92.5%; Score 393; DB 9; Length 2136;
Best Local Similarity 100.0%; Pred. No. 1e-113;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 33 GGATCGCTGCTCAAGAGGCGTCCGAGGAGAGAGGCCCAATCCTGTACTGAGACCTGG 92
DB 943 GGATCGCTGCTCAAGAGGCGTCCGAGGAGAGAGGCCCAATCCTGTACTGAGACCTGG 1002
QY 93 GAGCCCTCCGAAGGCCCATCATCTTCCCTGACTTGTGTAGAGCACTGTACCATTTCT 152
DB 1003 GAGCCCTCCGAAGGCCCATCATCTTCCCTGACTTGTGTAGAGCACTGTACCATTTCT 1062
QY 153 GGAGATGTTTCCCAATATCCACTGGGCTCCCGACGCCCACTTTTGAAGCAGGGGTG 212
DB 1063 GGAGATGTTTCCCAATATCCACTGGGCTCCCGACGCCCACTTTTGAAGCAGGGGTG 1122
QY 213 CCGCAACACAGAGTCTCTGAGACTGTAGACAGGAGAGCCGACGTTGGAACCCGGGAGCAG 272
DB 1123 CCGCAACACAGAGTCTCTGAGACTGTAGACAGGAGAGCCGACGTTGGAACCCGGGAGCAG 1182
QY 273 AGCCAGGTGGCCACAGTCTGACCAATGACATTCATGTCACGGGCGGTATGACTATCACT 332
DB 1183 AGCCAGGTGGCCACAGTCTGACCAATGACATTCATGTCACGGGCGGTATGACTATCACT 1242
QY 333 GGCACATCTACATCTACATGAGACAGTGTGGGGGACACACGGGCTCTGAGACCTC 392
DB 1243 GGCACATCTACATCTACATGAGACAGTGTGGGGGACACACGGGCTCTGAGACCTC 1302
QY 393 CCAAGTACCCCGGACACTCCATACCCCATTTCC 425
DB 1303 CCAAGTACCCCGGACACTCCATACCCCATTTCC 1335

```

```

RESULT 6
US-09-962-436-262
; Sequence 262, Application US/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
; FILE REFERENCE: 689290-75
; CURRENT APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,082
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/234,924
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 262
; LENGTH: 2136
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-436-262

```

```

Query Match          92.5%; Score 393; DB 10; Length 2136;
Best Local Similarity 100.0%; Pred. No. 1e-113;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 33 GGATCGCTGCTCAAGAGGCGTCCGAGGAGAGAGGCCCAATCCTGTACTGAGACCTGG 92
DB 943 GGATCGCTGCTCAAGAGGCGTCCGAGGAGAGAGGCCCAATCCTGTACTGAGACCTGG 1002
QY 93 GAGCCCTCCGAAGGCCCATCATCTTCCCTGACTTGTGTAGAGCACTGTACCATTTCT 152
DB 1003 GAGCCCTCCGAAGGCCCATCATCTTCCCTGACTTGTGTAGAGCACTGTACCATTTCT 1062
QY 153 GGAGATGTTTCCCAATATCCACTGGGCTCCCGACGCCCACTTTTGAAGCAGGGGTG 212

```

```

Dd      1243 GGGACATCTACAATTCATAAGTGCACCACTACTGCGGGGAGCACCAGCGGTCTGGAGACTCT 1302
Oy      393 CCAGCTACCCCCGNAACTCCATACCCCATCCC 425
          |||||
Db      1303 CCAGCTACCCCCGNAACTCCATACCCCATCCC 1335

RESULT 8
US-09-867-701-5601
; Sequence 5601, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Hallocke, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09-867,701
; CURRENT FILING DATE: 2001-05-29

```

```

RESULT 8
US-09-867-701-5601
? Sequence 5601, Application US/09867701
? Patent No. US20020132237A1
GENERAL INFORMATION:
? APPLICANT: Aglate, Paul A.
? APPLICANT: Jones, Robert
? APPLICANT: Harlocker, Susan L.
? TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
? TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
? FILE REFERENCE: 210121.497
? CURRENT APPLICATION NUMBER: US/09/867,701
? CURRENT FILING DATE: 2001-05-29
? NUMBER OF SEQ ID NOS: 10912
? SOFTWARE: FASTSEQ for Windows Version 4.0
? SEQ ID NO 5601
? LENGTH: 281
? TYPE: DNA
? ORGANISM: Homo sapien
US-09-867-701-5601

Query Match          59.7%; Score 253.8; DB 10; Length 281;
Best Local Similarity 97.1%; Pred. No. 3.1e-70;
Matches 269; Conservative 0; Mismatches 7; Indels 1; Gaps 1

QY 144 CCCATTTCTGAGATGTTTCCCGATATCCACTGGGCTCCCGCAGCC-CAGTTTGGG 202
DB 5 CCACCTTCTGAGATGTTTCCCGATATCCACTGGGCTCCCGCAGCCCGCAGTTTGGG 64

QY 203 GGCAGGGGTCGCCACACAGCAGAGTCTCTGAGACCTGACAGGGAGCGCAGTTGGAAC 262
DB 65 GGCAGGGGTCGCCACACAGCAGAGTCTCTGAGACCTGACAGGGAGCGCGGTTGGAAC 124

QY 263 CGGGAGAGCAGGACGAGGTCGCCACAGGTACCAATGGCATTCATGTCAACGGCGGCTAT 322.
DB 125 CGGGAGAGCAGTCAAGTGGCCACCGGTACCAATGGCATTCATGTCAACGGCGGCTAT 184

QY 323 GACTATCACTGGCAACATTCATCTACAAATGAGACCACTACTGGGGGAGCACCAGGGTCC 382
DB 185 GACTATCACTGGCAACATTCATCTACAAATGAGACCACTACTGGGGGAGCACCAGGGTCA 244

```

```

RESULT 9
US-09-867-701-5666/c
; Sequence 5666, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: fastSeq for Windows Version 4.0
; SEQ ID NO 5666
; LENGTH: 289
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-5666

```

Query Match 59.1%; Score 251.2; DB 10; Length 289;
Best Local Similarity 98.5%; Pred. No. 2.1e-69;
Matches 264; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

OY 155 AGATGTTTCCCACTATCCATGGGCTCCCGCAGCCCAATTGGAGGAGGGTGGCC 214
|||||
DB 289 AGATGTTTCCCACTATCCATGGGCTCCCGCAGCCCAATTGGAGGAGGGTGGCC 230
|||||
OY 215 GCACACAGACA-GTCTCTGTGACTGACACAGGAGCCGAGTTGAAACCGGGAGACAGA 273
|||||
DB 229 GCACACAGACAGTGTCTCTGTGACTGACACAGGAGCCGCGGTTGAAACCGGGAGACAGA 170
|||||
OY 274 GCCAGTGGGGCCACGATACCAATGGCAATGTCATGTCACGGGGGCTATGATCACTG 333
|||||
DB 169 GCCAGTGGGGCCACGATACCAATGGCAATGTCATGTCACGGGGGCTATGATCACTG 110
|||||
OY 334 GCACATCTACATCTACATGACAGACAGTACTGAGGGGAGCACCGGCTCTGAGACCTCC 393
|||||
DB 109 GCACATCTACATCTACATGACAGACAGTACTGAGGGGAGCACCGGCTCTGAGACCTCC 50
|||||
OY 394 CAGCTACCCCGAAGCTCCATCCCAT 421
|||||
DB 49 CAGCTACCCCGAAGCTCCATCCCAT 22
|||||

RESULT 10

US-09-907-372-16
; Sequence 16, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Warren, Bridget A.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 16
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 702245091H1
US-09-907-372-16

Query Match 40.4%; Score 171.8; DB 10; Length 371;
Best Local Similarity 81.0%; Pred. No. 1.8e-44;
Matches 200; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

OY 178 GGCTCCCGCAGCCCAATTGGAGCAGGGGTGCCGCAACAGCAGAGTCTCTGAGCC 237
|||||
DB 4 GACTCCAGACACCCCAAGTCTGAGAGAGATGTACACAGACAGAGTCTCTGAGCC 63
|||||
OY 238 TGACACAGGAGCCCAATGGAAACCGGGAGCAGAGCCAGAGTCCCGTCAATG 297
|||||
DB 64 AGGCCAGAGAGCTGAGCTCCAGAAAGGCCAGAGTCCCGTCAATG 123
|||||
OY 298 GCATTCAATGTCACGGGGGCTCTATGACTATCACTGGCAACATCTACATGAGAC 357
|||||
DB 124 GCATTCAATGTCACGGGGGCTCTATGACTATCACTGGCAACATCTACATGAGAC 183
|||||
OY 358 CAGTACTGGGGGAGCACCGGGTCTTGAGAGACTCCAGTACCCCGCAATCAATAC 417
|||||
DB 184 CAGTACTGGGGGAGCACCGGGTCTTGAGAGACTCCAGTACCCCGCAATCAATAC 243
|||||
OY 418 CCATTCC 424
|||||
DB 244 CCACCC 250
|||||

RESULT 11

US-09-907-372-13
; Sequence 13, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Warren, Bridget A.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 13
; LENGTH: 206
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 700302531H1
US-09-907-372-13

Query Match 32.0%; Score 135.8; DB 10; Length 206;
Best Local Similarity 81.0%; Pred. No. 3.1e-33;
Matches 158; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

OY 231 CTGACCTGACACGAGACCGCAGTGGAAACCCGGGGAGCAGACCCAGTGGCCACGGT 290
|||||
DB 2 CTGATCCAGGCGCAGAGAGCTGAGCGCTGGGGCAATGAGTGGCCACGGT 61
|||||
OY 291 ACCAATGCAATCAATGTCACCGGGGCTCTATGACTATCACTGCAACATCTACATGAC 350
|||||
DB 62 GCGAATGCAATCAATGTCACCGGGGCTCTATGACTATCACTGCAACATCTACATGAC 121
|||||
OY 351 AATGACAGTACTGAGGGGAGCACCGGGTCTTGAGAGACTCCAGTACCCCGAAGCT 410
|||||
DB 122 AATGGGCGAGTCTGAGGGGAGAACCGGGGCTTGAGAGACTCCAGTACCCCGAAGCT 181
|||||
OY 411 CCATACCCCATTCCTCC 425
|||||
DB 182 CCATACCCCATTCCTCC 196
|||||

RESULT 12

US-09-907-372-10
; Sequence 10, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Warren, Bridget A.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 10
; LENGTH: 219
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 697459H1
; LOCATION: 76, 131
; OTHER INFORMATION: a, t, c, g, or other
US-09-907-372-10

Query Match 23.2%; Score 98.4; DB 10; Length 219;
Best Local Similarity 98.0%; Pred. No. 1.7e-21;
Matches 99; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 325 CTATCACTGGCAACATCTACATGACAGCAGTACTGGGGGAGCACCGGGTCTG 384

Db 9 CGATCTACTGGCAACATCTACATCTACATGACAGACAGTACTGGGGGAGACCGGGTCTG 68
 QY 385 GAGACCTCCAGCTACACCCCGAAGCTTCATACCCCATTTCCC 425
 |||||||
 Db 69 GAGACCTNCCAGCTACCCCGAAGCTTCATACCCCATTTCCC 109

RESULT 13

US-09-907-372-11
 ; Sequence 11, Application US/09907372
 ; Patent No. US20020068242A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lal, Preeti G.
 ; APPLICANT: Warren, Bridget A.
 ; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
 ; FILE REFERENCE: PC-0050 US
 ; CURRENT APPLICATION NUMBER: US/09/907,372
 ; CURRENT FILING DATE: 2001-07-27
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 11
 ; LENGTH: 279
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; OTHER INFORMATION: Incyte ID No. US20020068242A1 3321983H1
 ; US-09-907-372-11

Query Match

Best Local Similarity 16.7%; Score 71; DB 10; Length 279;
 Best Local Similarity 100.0%; Pred. No. 7,2e-13;
 Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 355 GACCACTACTGGGGGAGACACCGGTCCTGGAGACCTCCAGTACCCCGAAGCTTCAT 414
 |||||||
 Db 1 GACCACTACTGGGGGAGACACCGGTCCTGGAGACCTCCAGTACCCCGAAGCTTCAT 60
 QY 415 ACCCATTTCCC 425
 |||||||
 Db 61 ACCCATTTCCC 71

RESULT 14

US-09-907-372-12/c
 ; Sequence 12, Application US/09907372
 ; Patent No. US20020068242A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lal, Preeti G.
 ; APPLICANT: Warren, Bridget A.
 ; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
 ; FILE REFERENCE: PC-0050 US
 ; CURRENT APPLICATION NUMBER: US/09/907,372
 ; CURRENT FILING DATE: 2001-07-27
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 12
 ; LENGTH: 862
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; OTHER INFORMATION: Incyte ID No. US20020068242A1 8576918T1
 ; US-09-907-372-12

Query Match

Best Local Similarity 11.2%; Score 47.8; DB 10; Length 862;
 Best Local Similarity 74.5%; Pred. No. 1.8e-05;
 Matches 114; Conservative 0; Mismatches 32; Indels 7; Gaps 4;

QY 279 GTGGCCGACGATGCAATGCTATGTCACCGCGGCTCTATGACTATCACTGGCAAC 338
 |||||||
 Db 853 GTGGCCGCGGTACCCAGGCAATAGGT-ACCGCGGTTCAAGACAATACACGGGCAAA 795

QY 339 A--TCTACATCTACATGAGCA--GTACTGGGGGAGACCGGGTCTCTGAGACCTTCC 394
 |||||||
 Db 794 AGTCGAATTCACACAGGAGCAAGTACTGGGGAGACACCGGGTCTCTGAGACCTTCC 735
 QY 395 AGCTACCCCGAAGC--TCCATACCCCATTTCCC 425
 |||||||
 Db 734 AGCTACCCCGAAGCCTTCATACCCCATTTCCC 702

RESULT 15

US-09-844-864-22
 ; Sequence 22, Application US/09844864
 ; Patent No. US20020042926A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Matzuk, Martin
 ; APPLICANT: Ren, Yongsheung
 ; APPLICANT: Wu, Xuemei
 ; TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
 ; FILE REFERENCE: P01925052 / 09807797 / OIA 99-48
 ; CURRENT APPLICATION NUMBER: US/09/844,864
 ; CURRENT FILING DATE: 2001-04-27
 ; PRIOR APPLICATION NUMBER: 60/106,020
 ; PRIOR FILING DATE: 1998-10-28
 ; PRIOR APPLICATION NUMBER: PCT/US99/25209
 ; PRIOR FILING DATE: 1999-10-28
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 22
 ; LENGTH: 809
 ; TYPE: DNA
 ; ORGANISM: mus musculus
 ; US-09-844-864-22

Query Match 8.3%; Score 35.4; DB 10; Length 809;
 Best Local Similarity 49.2%; Pred. No. 0.13;
 Matches 93; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 200 GGAAGCAGGGGTGCGCGACACAGAGTCTCTGAGCTGACAGAGAGCCGAGTTGA 259
 |||||||
 Db 610 GCAGCGCGGGGTGGAGACAGCCACACCGAGAGACCGGAGACGTGTGGGGGATGA 669
 QY 260 ACCCGGGGAGAGAGCCAGTGGCCACGCTACCAATGGCATTCATGTCACCGCGGGTC 319
 |||||||
 Db 670 GTCTGAGCCTGAG 729
 QY 320 TATGACTATCACTGCGACATCTACTATCTACATGAGACAGTACTGGGGAGCCAGCGG 379
 |||||||
 Db 730 TTCGATGACCTCTCCCGCAAGACACCATACCTTGAGCAGAGCAAGAGCTCTCGGTT 789
 QY 380 TCCTGGAGA 388
 |||||||
 Db 790 TCCAGGTGA 798

Search completed: April 15, 2003, 18:45:54
 Job time : 33.6131 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2003, 10:39:00 ; Search time 100.94 Seconds
(without alignments)
526.717 Million cell updates/sec

Title: US-09-917-372-1
Perfect score: 2248
Sequence: 1 MLLPWATAPGLAWPLVLG.....EHCATPSNRGPRNQFIETHD 399

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :			
A_Geneseq_101002.*			
1:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*	415	349.5
2:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*	415	349.5
3:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*	415	349.5
4:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*	415	349.5
5:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*	415	349.5
6:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*	415	349.5
7:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*	415	349.5
8:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*	415	349.5
9:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*	415	349.5
10:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*	415	349.5
11:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*	415	349.5
12:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*	415	349.5
13:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*	415	349.5
14:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*	415	349.5
15:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*	415	349.5
16:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*	415	349.5
17:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*	415	349.5
18:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*	415	349.5
19:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*	415	349.5
20:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*	415	349.5
21:	/SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*	415	349.5
22:	/SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*	415	349.5
23:	/SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*	415	349.5

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1404	62.5	415	AAB36700	Human tumour necro
2	1325.5	59.0	305	ABP41926	Human ovarian anti
3	1129	50.2	197	AAW23220	Extracellular doma
4	1129	50.2	197	AAW31326	Human lymphotoxin
5	456	20.3	77	AAW94642	TNF-R extracellular
6	456	20.3	77	AAW69194	Human TNF-R extrac
7	363	16.1	461	AAW22504	p75 Tumour Necrosi
8	351.5	15.6	518	AAW51003	Sequence of a reco
9	349.5	15.5	461	AAW11001	40kD TNF inhibitor
10	349.5	15.5	461	AAW11141	Human TNF-R deduce

11	349.5	15.5	461	14	AAW42058	Fibroblast derived
12	349.5	15.5	461	21	AAB37801	Human tumour necro
13	349.5	15.5	461	21	AAB18717	A human tumour nec
14	349.5	15.5	461	21	AAB01342	Death receptor. H
15	349.5	15.5	461	22	AAB35331	Human TNF receptor
16	349.5	15.5	461	22	AAB36698	Human tumour necro
17	349.5	15.5	461	22	AAB37686	Human 40 kDa TNF i
18	349.5	15.5	461	23	AAU75172	Human TNF receptor
19	349.5	15.5	461	23	AAU75173	Human TNF receptor
20	349.5	15.5	461	23	AAU75174	Human TNF receptor
21	339.5	15.1	518	22	AAB70001	STNFR(075):Fc fusi
22	339.5	15.1	518	22	AAB50080	TNFR:Fc fusion pro
23	337.5	15.0	485	13	AAW24016	Fusion protein TNF
24	330	14.7	487	22	AAB50084	TNFR:Fc fusion pro
25	325	14.5	474	21	AAW94718	Human type 2 tumou
26	322.5	14.3	474	12	AAW11142	TNF-R deduced from
27	321.5	14.3	225	21	AAW77463	Primate protein se
28	321.5	14.3	227	22	AAB66981	Tnfr2 protein. Un
29	319	14.2	300	19	AAW66102	Amino acid sequenc
30	319	14.2	300	19	AAW63622	Human tumour necro
31	319	14.2	300	20	AAW03099	Human lung TNF-rec
32	319	14.2	300	20	AAW42182	Human FLINT #1 pro
33	319	14.2	300	20	AAW17479	Mammalian tumour n
34	319	14.2	300	20	AAW06817	Human Dcr3 polyep
35	319	14.2	300	20	AAW97749	Human tumour necro
36	319	14.2	300	20	AAW95082	Orphan receptor (H
37	319	14.2	300	21	AAW19335	A full length huma
38	319	14.2	300	21	AAW28559	Human soluble TNF
39	319	14.2	300	21	AAW24057	Human PRO212 prote
40	319	14.2	300	21	AAW33416	Human PRO212 prote
41	319	14.2	300	21	AAW03621	Human Fas ligand i
42	319	14.2	300	21	AAW97246	M68 TNF receptor i
43	319	14.2	300	21	AAW90357	Human tumour necro
44	319	14.2	300	21	AAW24395	Human PRO212 prote
45	319	14.2	300	21	AAW96596	Human FLINT. Homo

ALIGNMENTS

RESULT 1
AAB36700
ID AAB36700 standard; Protein; 415 AA.
XX AC AAB36700;
XX DT 15-MAR-2001 (first entry)
XX DE Human tumour necrosis factor receptor LTBR protein SEQ ID NO:6.
XX KW Human; tumour necrosis factor receptor 5; TRID; TNFR-5; TR5; nontropic;
XX KW TRAIL receptor without intracellular domain; diagnosis; cytostatic;
XX KW tumour necrosis factor related apoptosis inducing ligand; vasotropic;
XX KW immunosuppressive; neuroprotective; antiviral; antiinflammatory;
XX KW anticonvulsant; antiparasitic; cardiant; anti-HIV; antiparkinsonian;
XX KW gene therapy; restenosis; graft versus host disease; tumour; cancer;
XX KW apoptotic cell death related disease; autoimmune disorder;
XX KW cardiovascular disorder; viral infection.
XX OS Homo sapiens.
XX PN WO200071150-A1.
XX PD 30-NOV-2000.
XX PF 18-MAY-2000; 2000WO-US13515.
XX PR 20-MAY-1999; 99US-0135164.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Wei Y, Ruben SM, Gentz RL, Ni J;
XX


```
CC selectively or partially block the LT-beta-R pathway may be useful
CC in the treatment of abnormal lymphoid organ development associated
CC with misexpression or overexpression of signalling by the LT-beta-R
CC pathway. The present LT-beta-R blocking agent is capable of
CC selectively inhibiting Th1, but not Th2 cell dependent immune
CC effector mechanisms. As Th1 cytokines can inhibit Th2 cell
CC dependent responses, the present LT-beta-R blocking agent may also
CC indirectly stimulate certain Th2 cell dependent responses which are
CC normally inhibited by Th1 induced cytokines. Doses of about 1 mg/kg
CC of the present soluble LT-beta-R are expected to be suitable
CC starting doses for optimising treatment.
XX
SQ Sequence 197 AA;

Query Match          50.2%; Score 1129; DB 18; Length 197;
Best Local Similarity 99.5%; Pred. No. 2.5e-67;
Matches 196; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 28 SQQAVPPYASENQTCDQOEKEYEPOHRIICSCRCPGTYVYSAKSRIRDVFVCATCAENS 87
Db | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
   1 SQQAVPPYASENQTCDQOEKEYEPOHRIICSCRCPGTYVYSAKSRIRDVFVCATCAENS 60

QY 88 YNEHWNYLTICQLCRPCDPVMGLEEIIAPCTSKRKTQCQPQMFCAAWALECTHCELLSD 147
Db | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
   61 YNEHWNYLTICQLCRPCDPVMGLEEIIAPCTSKRKTQCQPQMFCAAWALECTHCELLSD 120

QY 148 CPPGTEAELKDVEKGNNHCVPCKAGHFQNTSSPSARCQPHTRCENQGLVEAAPGTQAQSD 207
Db | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
   121 CPPGTEAELKDVEKGNNHCVPCKAGHFQNTSSPSARCQPHTRCENQGLVEAAPGTQAQSD 180

QY 208 TTCKNPLEPLPPEMSGSGS 224
Db | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
   181 TTCKNPLEPLPPEMSGSGT 197

RESULT 4
AAY31326
ID AAY31326 standard; peptide; 197 AA.
XX
AC AAY31326;
XX
DT 04-OCT-1999 (first entry)
XX
DE Human lymphotoxin (LT)beta-receptor extracellular region.
XX
KW Lymphotoxin-beta; LT-beta; LT-beta receptor; follicular dendritic cell;
KW immune system; tumour; follicular lymphoma; extracellular domain; human.
XX
OS Homo sapiens.
XX
PN WO9938525-A1.
XX
PD 05-AUG-1999.
XX
PF 29-JAN-1999; 99WO-US01928.
XX
PR 02-FEB-1998; 98US-0073410.
XX
PR 30-JAN-1998; 98US-0073112.
XX
PA (BIOJ ) BIOGEN INC.
XX
PI Browning J, Thorbecke J, Tsiagbe V;
XX
DR WPI; 1999-469242/39.
XX
PT New method of treating follicular lymphomas by inhibiting
XX interaction between lymphotoxin-beta and its receptor
XX
PS Example 1; Page 25-26; 31pp; English.
XX
CC The invention provides a method for arresting or reducing, severity of
CC effects of a tumour by administration of a composition which inhibits
CC the interaction between lymphotoxin (LT)-beta and its receptor. An
```

CC inhibitor of the interaction between LT-beta and its receptor can be
 CC administered for altering the survival or maintenance of follicular
 CC dendritic cells in a subject and for altering the architecture of the
 CC organs of the immune system. The method is useful for treating tumours,
 CC specifically follicular lymphomas. It offers an alternative therapy for
 CC those with tumours resistant to traditional chemotherapy. The present
 CC sequence represents the extracellular region of the human LTbeta-receptor
 CC and comprises the ligand binding domain.

SQ Sequence 197 AA;

Query Match 50.2%; Score 1129; DB 20; Length 197;
 Best Local Similarity 99.5%; Pred. No. 2.5e-67;
 Matches 196; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 28 SQQAVPPYASENQCRDQKEYEYEPQHRICCSRCPPGTYYSAKCSRIKDTVCATCAENS 87
 DB 1 SQQAVPPYASENQCRDQKEYEYEPQHRICCSRCPPGTYYSAKCSRIKDTVCATCAENS 60
 QY 88 YNEHNNYLTICQLCRPCDPVGMGLEIAPCTSKRKTQCRCPGMFCAAWALECTHCELLSD 147
 DB 61 YNEHNNYLTICQLCRPCDPVGMGLEIAPCTSKRKTQCRCPGMFCAAWALECTHCELLSD 120
 QY 148 CPPGTEAEKDEVGKGNHCVCKAGHFQNTSSPSARCQPHTRCENQGLVEAAPGTAQSD 207
 DB 121 CPPGTEAEKDEVGKGNHCVCKAGHFQNTSSPSARCQPHTRCENQGLVEAAPGTAQSD 180
 QY 208 TTCKNPLEPLPPMSGS 224
 DB 181 TTCKNPLEPLPPMSGT 197

RESULT 5

ID AAW94642 standard; peptide; 77 AA.

AC AAW94642;

DT 29-APR-1999 (first entry)

DE TNF-R extracellular Cys-rich domain TNF-R-tp.

XX Tumour necrosis factor receptor; TNF-R; autoimmune diseases;
 KW inflammation; septic shock; cachexia; graft versus host disease;
 KW skin allergic reaction; immune complex disease; malaria;
 KW transplantation rejection.

OS Homo sapiens.

XX WO9853842-A1.

PN 03-DEC-1998.

PD 29-MAY-1998; 98WO-US10891.

PF 30-MAY-1997; 97US-0866545.

XX (UYPE-) UNIV PENNSYLVANIA.

PI Greene MI, Murali R, Takasaki W;

DR WPI; 1999-080781/07.

XX New compounds designed from a binding loop of a tumour necrosis
 PT factor receptor - are capable of inhibiting the biological
 PT activities of tumour necrosis factor, e.g., in treating inflammation
 PT or autoimmune diseases

PS Disclosure; Fig 1; 78pp; English.

XX The present invention describes peptides and peptide analogues which
 CC correspond in primary sequence to a binding loop of a tumour necrosis
 CC factor receptor (TNF-R) superfamily member. The compounds are especially

CC designed from a binding loop of TNF-R p55. They are capable of
 CC inhibiting TNF binding to its cellular receptors and may be used to
 CC inhibit the biological activities of TNF. They may be used in treating
 CC TNF-associated conditions such as acute and chronic inflammatory
 CC responses, septic shock, cachexia, autoimmunity, graft-versus-host
 CC disease, skin allergic reactions, immune complex disease,
 CC transplantation rejection and malaria. Administration is, e.g. oral,
 CC transdermal, transmucosal, pulmonary, subcutaneous, intravenous or
 CC intramuscular. Parenteral dosage is 0.1-5 mg/kg/day. The present
 CC sequence represents an extracellular Cys-rich domain of TNF-R from the
 CC present invention.

SQ Sequence 77 AA;

Query Match 20.3%; Score 456; DB 20; Length 77;

Best Local Similarity 100.0%; Pred. No. 3.2e-23;

Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 VCATCAENSYNHNNYLTICQLCRPCDPVGMGLEIAPCTSKRKTQCRCPGMFCAAWALE 138

DB 1 VCATCAENSYNHNNYLTICQLCRPCDPVGMGLEIAPCTSKRKTQCRCPGMFCAAWALE 60

QY 139 CTHCELLSDCPPGTEAE 155

DB 61 CTHCELLSDCPPGTEAE 77

RESULT 6

ID AAB69194 standard; protein; 77 AA.

AC AAB69194;

DT 30-APR-2001 (first entry)

DE Human TNF-R extracellular Cys-rich domain TNF-R-tp SEQ ID NO:3.

XX Tumour necrosis factor receptor; TNF-R; inhibition; osteoclast;
 KW osteoclastogenesis; bone loss; bone resorption; osteopathic; cytostatic;
 KW antirheumatic; antiarthritic; antiinflammatory; immunomodulatory;
 KW tumour necrosis factor-related activation-induced cytokine; TRANCE;
 KW receptor activator of NF-kappaB ligand; RANK; osteoporosis;
 KW Paget's disease; metastatic bone disease; rheumatoid arthritis;
 KW periodontal disease; modulating dendritic cell maturation;
 KW T cell proliferation; CD40 receptor system.

OS Homo sapiens.

XX WO200108699-A1.

PN 08-FEB-2001.

PD 28-JUL-2000; 2000WO-US20510.

PF 28-JUL-1999; 99US-0146090.

PR (UYPE-) UNIV PENNSYLVANIA.

PA (AOKI/) AOKI K.

PA (HORN/) HORNE W C.

PA (BARO/) BARON R.

XX Aoki K, Horne WC, Baron R, Greene MI, Murali R;

PI WPI; 2001-182866/18.

XX Use of peptides and peptide analogs which are TRANCE/RANK inhibitors,
 PT for inhibiting osteoclastogenesis and bone resorption

PS Disclosure; Fig 1; 81pp; English.

XX The present invention describes a method for inhibiting
 CC osteoclastogenesis and bone resorption. Osteoclastogenesis and bone
 CC resorption inhibiting peptide analogues from the present invention have

PT esp. soluble form of TNF receptor, opt. as fusion protein with
 PT human immunoglobulin Fc region, esp. for treating arthritis
 XX
 PS Disclosure; Page 32-34; 47pp; English.

XX AAQ45224 is cDNA from clone 1 of library WI-26 V44 of human fibroblast
 CC cell line WI-26 V44. The mature full-length TNFRI is a glycoprotein
 CC having a mol. wt. of about 75-80 kDa. The cloning of the cDNA for
 CC TNFRI was described in Smith et al., Science 248:1019, 1990. Clone 1
 CC is contained in expression vector pCAV/NOT-TNFR (ATCC 68088). A
 CC recombinant chimeric antibody may be produced having TNFR sequences
 CC substituted for the variable domains of either or both of the
 CC immunoglobulin molecule heavy and light chains and having
 CC unmodified constant region domains. A specific example of a TNFR/Fc
 CC fusion protein is given in AAQ45225/R51003. The rhu TNFR-Fc fusion
 CC gene was created by ligating the following fragments into a cloning
 CC vector: 1) an 867 bp Asp718-Pvu2 fragment from pCAV/NOT-TNFR (ATCC
 CC 68088) contg. the cDNA encoding the truncated TNFR. 2) a 700 bp
 CC Styl-Spel fragment from plasmid pIXY498 coding for 232 AAs of the
 CC Fc portion of human IgG1. Plasmid pIXY498 is a yeast expression
 CC vector contg. the Fc fragment of human IgG1. 3) An oligo linker, to
 CC fuse the truncated TNFR with the human IgG1 Fc fragment. This
 CC linker was created by PCR using primer AAQ45226, which encodes the 3'
 CC end of the truncated TNF receptor and the 5' end of human IgG1,
 CC and primer AAQ45227, which is an antisense sequence encoding bps
 CC 257-237 of human IgG1.

XX Sequence 518 AA;

Query Match 15.6%; Score 351.5; DB 15; Length 518;
 Best Local Similarity 27.6%; Pred. No. 2.3e-15;
 Matches 106; Conservative 32; Mismatches 127; Indels 119; Gaps 16;

Qy 4 PWATSAAGLPLVLGI---FGLIAASQPOAVPYASE-NOTCRDQKEYEYPOHRIICC 59
 Db 26 PPASMAPVAVWAALAVGLWAGHALPAQVAFTPYAPPGSTCR--LREYDQTAQMCCK 83
 Qy 60 SRCPPGTYSACSRIRDTVCATCAENSYNEHWNLTICOLCR---PCDPVMGLEEIAPC 116
 Db 84 SKCSPQAHKVFCTKTSDTVCDSCEDSTYTLWNWVPECLSCGSRCSQDV---ETQAC 139
 Qy 117 TSKRTQRCQPGMFCMAWALE-CTHCELLSDCPPG-----TEAEKDEVGKGNHCVPC 170
 Db 140 TREQNRICTRPGWYCALSKQEGCRLCAPLKRCPGFGVARGTETSDVW-----CKPC 193
 Qy 171 KAGHPONTSSPSARCOPHTRCENOGIVEAAPGTAQSDTCKNPLEPLPPMSSGSLKKRP 230
 Db 194 APGTSNTSSDTCRPHQICN-----VVAIPGNASMDAVCTIS-----231
 Qy 231 QGEGNPVAGSWEPKPAHYPDLVQPLLPISGDVSPYSTGLPAAPVLEAGVPQQSPLD 290
 Db 232 -----TSPTSRMAPGA-----VHLPPQVPS---250
 Qy 291 LTREPQLEBGEQSOVAHGNTGHTVGGSMITGNIYIYINGPVLGPPGPGDLPATPE--- 347
 Db 251 -TRSQHTQTPPTSTSTSLFLPMGPS-----PPAEGSTGDEPKSCD 292
 Qy 348 -----PPYPIPE-EGDPG-----pp 361
 Db 293 KTHTCPCPAPELGPGSVFLFPP 316

RESULT 9

AAAR11001
 ID AAR11001 standard; Protein; 461 AA.

XX AAR11001;

XX 13-MAY-1991 (first entry)

XX 40kD TNF inhibitor precursor.

XX Tumour necrosis factor; inhibitor.

XX Homo sapiens.
 OS
 XX AU0958976-A.
 PN
 XX 24-JAN-1991.
 PD
 XX 16-JUL-1990; 90AU-0058976.
 PF
 XX 07-FEB-1990; 90US-0479661.
 PR
 XX 18-JUL-1989; 89US-0381080.
 PR
 XX 11-DEC-1989; 89US-0450329.
 XX
 PA (SYNE-) SYNERGEN INC.
 XX
 XX WPI; 1991-073847/11.
 DR
 XX N-PSDB; AAQ10907.

XX Tumour necrosis factor inhibitor - for suppression of TNF-alpha
 PT and -beta, useful as therapeutic agent.
 XX
 PS Disclosure; Fig 39; 142pp; English.
 XX
 CC The sequence comprises the entire 40 kD TNF inhibitor. The clone
 CC from which the sequence was deduced was isolated from a cDNA
 CC library prepd. from RNA form U937 cells treated with PMA/PHA.
 CC The whole gene can be inserted into expression vectors for prepn.
 CC of TNF inhibitor for use in the treatment of inflammatory and
 CC degenerative diseases.

CC See also AAR10986 and AAR10984.

XX Sequence 461 AA;

Query Match 15.5%; Score 349.5; DB 12; Length 461;
 Best Local Similarity 29.7%; Pred. No. 2.7e-15;
 Matches 113; Conservative 47; Mismatches 143; Indels 77; Gaps 22;

Qy 9 APLGAWPLVLGLFLLAASQPOAVP-----PYASE-NOTCRDQKEYEYPOHRIICSR 61
 Db 2 APVAVWAALAVGL-ELWAAA--HALPAQVAFTPYAPPGSTCR--LREYDQTAQMCCK 56
 Qy 62 CPPGTYSACSRIRDTVCATCAENSYNEHWNLTICOLCR---PCDPVMGLEEIAPCTS 118
 Db 57 CSPQAHKVFCTKTSDTVCDSCEDSTYTLWNWVPECLSCGSRCSQDV---ETQACTR 112
 Qy 119 KRKTQRCQPGMFCMAWALE-CTHCELLSDCPPG-----TEAEKDEVGKGNHCVPC 172
 Db 113 EQNRICTRPGWYCALSKQEGCRLCAPLKRCPGFGVARGTETSDVW-----CKPCAP 166
 Qy 173 GHFQNTSSPSARCOPHTRCENOGIVEAAPGTAQSDTTC--KNPLEPLPPMSSGSL----- 225
 Db 167 GTFSTNTSSDTCRPHQICN-----VVAIPGNASMDAVCTSTPTSRMAP---GAVHLPQ 219
 Qy 226 LKRPOGEGNPVAGSWEPKPAH-----PYFPDLVQPLLPISGDVS-PVS--TGLPAAP 276
 Db 220 VTRSQHTQTPPTSTSTSLFLPMGPS---PPAEGSTGDFALPVLGVLIVGVTRALG 272
 Qy 277 VLEAGVPQ-----QQSPDLTREPQLE--PGEQSQVAHGNTGHTV-----TGGSM 320
 Db 273 LLIGVNCVIMTQVKKKPLCLQREAKVPHLPADKARTQGEQHQHLLITAPSSSSSLE 332
 Qy 321 ITGNIYIYINGPVLGPPGPG 340
 Db 333 SSASALDRRAPTRNQPOAPG 352

RESULT 10

AAAR11141
 ID AAR11141 standard; Protein; 461 AA.

XX AAR11141;

XX 24-MAY-1991 (first entry)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2003, 10:45:16 ; Search time 102.845 Seconds
(without alignments)
799.386 Million cell updates/sec

Title: US-09-917-372-1
Perfect score: 2248
Sequence: 1 MLLPWATSAPLGLVGLG.....EHCGATPSNRGPRNQFIETHD 399

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

- 1: sp.archaea.*
- 2: sp.bacteria.*
- 3: sp.fungi.*
- 4: sp.human.*
- 5: sp.invertebrate.*
- 6: sp.mammal.*
- 7: sp.mhc.*
- 8: sp.organelle.*
- 9: sp.phase.*
- 10: sp.plant.*
- 11: sp.prodent.*
- 12: sp.virus.*
- 13: sp.vertebrate.*
- 14: sp.unclassified.*
- 15: sp.rvirus.*
- 16: sp.bacteriap.*
- 17: sp.archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	340	15.1	433	11 Q912M6	Q912M6 rattus norv
2	312	13.9	459	11 Q62327	Q62327 mus musculus
3	311.5	13.9	482	11 Q88734	Q88734 mus musculus
4	303.5	13.5	278	6 Q8SQ34	Q8SQ34 sus scrofa
5	271	12.1	277	6 Q8WMO2	Q8WMO2 ovis aries
6	266	11.8	651	13 Q98SM6	Q98SM6 gallus gall
7	253	11.3	276	13 Q9DD22	Q9DD22 gallus gall
8	252.5	11.2	267	6 Q02764	Q02764 oryctolagus
9	248	11.0	186	12 Q72735	Q72735 cowpox viru
10	244	10.9	283	6 Q9XSZ8	Q9XSZ8 cercopithec
11	242.5	10.8	302	13 Q9PUS0	Q9PUS0 salvelinus
12	236	10.5	285	13 Q9OW71	Q9OW71 oncorhynch
13	233	10.4	186	12 Q9FP87	Q9FP87 cowpox viru
14	230	10.2	351	12 Q75117	Q75117 cowpox viru
15	228	10.1	285	13 Q90YS6	Q90YS6 oncorhynch
16	227	10.1	186	12 Q911R5	Q911R5 vaccinia vi

17	226.5	10.1	167	12 Q8UYL3	Q8UYL3 vaccinia vi
18	225	10.0	349	12 Q57098	Q57098 camelpox vi
19	225	10.0	349	12 Q57099	Q57099 monkeypox v
20	225	10.0	349	12 Q8UYA7	Q8UYA7 camelpox vi
21	225	10.0	349	12 Q57284	Q57284 camelpox vi
22	223.5	9.9	446	6 Q95ND3	Q95ND3 felis silve
23	222.5	9.9	169	11 Q9JKE0	Q9JKE0 rattus norv
24	222	9.9	186	12 Q9WJB4	Q9WJB4 vaccinia vi
25	221	9.8	349	12 Q57097	Q57097 camelpox vi
26	221	9.8	349	12 Q57100	Q57100 monkeypox v
27	219	9.7	349	12 Q57102	Q57102 monkeypox v
28	219	9.7	349	12 Q57291	Q57291 monkeypox v
29	218.5	9.7	167	12 Q9DJL2	Q9DJL2 cowpox viru
30	218.5	9.7	350	12 Q57116	Q57116 cowpox viru
31	218	9.7	348	12 Q57103	Q57103 monkeypox v
32	218	9.7	348	12 Q57108	Q57108 monkeypox v
33	216.5	9.6	167	12 Q72762	Q72762 cowpox viru
34	216	9.6	348	12 Q57277	Q57277 monkeypox v
35	216	9.6	349	12 Q57101	Q57101 monkeypox v
36	210	9.3	348	12 Q57112	Q57112 variola vir
37	210	9.3	348	12 Q85407	Q85407 variola vir
38	208	9.3	351	12 Q57121	Q57121 cowpox viru
39	207	9.2	349	12 Q57110	Q57110 variola vir
40	207	9.2	349	12 Q57111	Q57111 variola vir
41	207	9.2	349	12 Q89098	Q89098 variola vir
42	207	9.2	349	12 Q89118	Q89118 variola vir
43	207	9.2	351	12 Q73559	Q73559 cowpox viru
44	205	9.1	347	12 Q57119	Q57119 cowpox viru
45	203.5	9.1	328	6 Q9BDP0	Q9BDP0 aotus trivi

ALIGNMENTS

RESULT 1

Q912M6 PRELIMINARY; PRT; 433 AA.
ID Q912M6;
AC Q912M6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE Tumor necrosis factor receptor type II (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RA Osburg B., Peiser C., Doemling D., Schomburg L., Voigt K., Bickel U.;
RT "TNF-receptors p60 and p80 are constitutively expressed by rat brain
capillary endothelial cells and participate in TNF-alpha transport
through the blood-brain barrier.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF420214; AAL16021.1; -;
DR InterPro; IPR001368; TNFR_c6.
DR Pfam; PF00020; TNFR_c6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_2.
DR PROSITE; PSS0050; TNFR_NGFR_2; 3.
FT Receptor.
FT NON_TER 1
FT NON_TER 433
SQ SEQUENCE 433 AA; 45723 MW; 75736D835E72CA4A CRC64;

Query Match 15.1%; Score 340; DB 11; Length 433;

Best Local Similarity 30.0%; Pred. No. 2.9e-19;

Matches 104; Conservative 45; Mismatches 144; Indels 54; Gaps 19;

QY 35 PYASENCTCRDQKEYEYPOHRTCCSRCPPTVVSACSRIRDTVCATCAENSNEHWNY 94

DB 11 PYKPEPGNQCIQSQYDKKAQCCAKCPGPAKFKHKTSDTVCADCAAGMFTQVNNH 70

QY 95 LTICQLC-RPC--DPVNGLEIAPCTSKRKTCRCQPGMFC--AWALECTHCELLSDCP 149

```

Db 71 LHTCLSCSSCDDQV-----ETHNCTKKONRYCACNADSYCALKLHSGNCRQCMKLSKGG 126
QY 150 PCTEAEKDEKGVKNNHCVPCKAGHFONTSSPSARCOPHTRCENQGLVEAAPGTAQSDTT 209
Db 127 PGF-GVARSTNGNVICSAACAGTFSDDTSTDVCRPHRIS-----ILAIPEGNASTDAY 181
QY 210 CKNPLEPLPEMSSGSLKRRPQGGPNPVAGSWEP- PKAHYPFP-----DLVQPLLPITS 262
Db 182 CASE-SPTSPASVPTIYVSQPEPTRSQPM--DQEPGSPQTHIPVSLGSTPIIEP--SIT 236
QY 263 GDVS-PVS--TGLPAAPVLEAG-----VPOQSQPLDLTRE---PQLEPGSQSQVAHGT 309
Db 237 GGISLPIGLIVGLTGLGLMLGLANCFTLVQRKKPKSCLORETMPHL--PDEKSDQDAVGL 295
QY 310 NGIHV--TGGSMITIGNIYINGVPLGGPPGPGDLPATPEPPYPIPE 354
Db 296 EQOHLITAPSSSSSS-----LESSASAGDRRA-----PPGVIPQ 330

```

RESULT 2

```

Q62327
ID Q62327 PRELIMINARY; PRT; 459 AA.
AC Q62327
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Murine tumour necrosis factor receptor 2 protein (Fragment).
GN TNFRSF1B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN NCBI_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=NOD;
RA Powell E.E., Wicker L.S., Peterson L.B., Todd J.A.;
RT "Amino acid variation in the tumor necrosis factor receptor 2 is
RT linked to autoimmune diabetes in NOD mice.";
RL Genomics 0:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD;
RX MEDLINE=95178848; PubMed=7873884;
RA Powell E.E., Wicker L.S., Peterson L.B., Todd J.A.;
RT "Allelic variation of the type 2 tumor necrosis factor receptor
RT gene.";
RL Mamm. Genome 5:726-727(1994).
DR EMBL; X76401; CAA53981.1; -.
DR HSSP; P19438; INCF.
DR MGD; MGI:1314883; Tnfrsf1b.
DR InterPro; IPR001368; TNFR_c6.
DR Pfam; PF00020; TNFR_c6; 4.
DR SMART; SM00208; TNFR_4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 3.
KW Receptor.
FT NON_TER 1 1
FT VARIANT 87 87 S -> T.
FT VARIANT 93 93 T -> I.
FT VARIANT 268 268 F -> I.
FT VARIANT 345 345 S -> F.
FT VARIANT 421 421 Y -> C.
SQ SEQUENCE 459 AA; 48686 MW; 6C51D2CFC14626DF CRC64;

```

Query Match 13.9%; Score 312; DB 11; Length 459;

Best Local Similarity 29.28%; Pred. No. 5.5e-17;

Matches 89; Conservative 41; Mismatches 135; Indels 40; Gaps 13;

```

QY 35 PYASE-NQTCRQDEKYEYEPQHRICCSRCPPGTYYSAKCSRIKRDIVCATCAENSNEHWN 93
Db 16 PKPPEPGYECQISQ-EYDRAQOMCCAKCPGQYVYKHCNKTSITDVCADEASMTQWN 74

```

```

QY 94 YLTICQLCR---PCDPVMGLEEIAPICTSKRKTQCRQCFMCA--AWALECTHCELLSDC 148
Db 75 QFRTCLSCSSCSTQV---ETRACTKQONRVACAGRYCALKTHSGSCRCQMRLSKC 130
QY 149 PGTEAEKDEKGVKNNHCVPCKAGHFONTSSPSARCOPHTRCENQGLVEAAPGTAQSDT 208
Db 131 GFGF-GVASSRAPNGNVLCACAPGTFSDDTSTDVCRPHRIS-----ILAIPEGNASTD 185
QY 209 TCKNPLEPLPEMSSGSLKRRPQGGPNPVAGSWEP- PKAHYPFPDLPVQL----- 258
Db 186 VCA-PESTLTAIPRTLYVSOPEPTRSQPL--DQEPGFSQT--PSILISLSTPIIEOST 240
QY 259 -----LPISGDVSPYSTGLPAAPVLEAG--VPOQSQPLDLTREPOLE--PGSQSQVAHGT 309
Db 241 KGGISLPIGLIVGLTGLMLGLVNCFTLVQRKKPKSCLOQDAKVPVHPDDEKSDQAVGL 300
QY 310 NGIHV 314
Db 301 EQOHL 305

```

RESULT 3

```

O88734
ID O88734 PRELIMINARY; PRT; 482 AA.
AC O88734
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE P80 TNF-alpha receptor.
GN TNFR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN NCBI_TaxID=10090;
RP SEQUENCE FROM N.A.
RX MEDLINE=98414512; PubMed=9740674;
RA Hurlie B., Segade F., Rodriguez R., Ramos S.S., Lazo P.S.;
RT "The Mouse Tumor Necrosis Factor Receptor 2 Gene: Genomic Structure
RT and Characterization of the two Transcripts.";
RL Genomics 52:79-98(1998).
DR EMBL; Y14619; CAA74969.1; -.
DR EMBL; Y14620; CAA74969.1; JOINED.
DR EMBL; Y14621; CAA74969.1; JOINED.
DR EMBL; Y14622; CAA74969.1; JOINED.
DR EMBL; Y14623; CAA74969.1; JOINED.
DR EMBL; Y14679; CAA74969.1; JOINED.
DR HSSP; P19438; INCF.
DR InterPro; IPR001368; TNFR_c6.
DR Pfam; PF00020; TNFR_c6; 4.
DR SMART; SM00208; TNFR_4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 3.
KW Receptor.
SQ SEQUENCE 482 AA; 51106 MW; F6C15046B48FF83C CRC64;

```

Query Match 13.9%; Score 311.5; DB 11; Length 482;

Best Local Similarity 27.5%; Pred. No. 6.3e-17;

Matches 102; Conservative 48; Mismatches 162; Indels 59; Gaps 17;

```

QY 9 APGLAWGPLVLGLFGLLAASQFQAVP-----PYASE-NQTCRQDEKYEYEPQHRICCSR 61
Db 2 APAALWVALV---FELQLWATHTVPAQVLTVPKPEPGYECQISQ-EYDRAQOMCCAK 57
QY 62 CPPGTYYSAKCSRI-----RDTVCATCAENSNEHWNLTICOLCR---PCDPVMGLE 111
Db 58 CPPGQYVYKHCNKTSITDVCADEASMTQWNQFRTCLSCSSCSTQV----- 113
QY 112 EIAPCTSKRKTQCRQCFMCA--AWALECTHCELLSDCPPGTEAEKDEKGVKNNHCVP 169
Db 114 ETTRACTKQONRVACAGRYCALKTHSGSCRCQMRLSKCQPGF-GVASSRAPNGNVLC 172
QY 170 CKAGHFONTSSPSARCOPHTRCENQGLVEAAPGTAQSDTTCKNPLEPLPEMSSGSLKRR 229

```

```
Db 173 CAPGTFSDTSDVCRPHRCS-----ILAIIGNASTDAVCA-PESPTLSAIPRLTYVSQ 277
Qy 230 POGEGPNVAGSWEPKKAHPFDVLQPL-----LPISGDVSPVSTGLPA 274
Db 228 PEPTRSQPL--DOEFGPSQT--PSILTSLSGSPPIIEQSTKGGISLPIGLIVGVTSLGLLM 283
Qy 275 APVLEAG--VPOQSPDLDTREPOLE--PGEOSQVARGTNGJHVTGGSTTIGNIYVNG 330
Db 284 LGLVNCFLVQRKKFSCLODKAKVPHVPDEKSDAVGLEQOHLTTAPSSSSSSSSLESS 343
Qy 331 PVLG---GPPG 338
Db 344 ASAGDRRAPPG 354

RESULT 4
Q8SQ34
ID Q8SQ34 PRELIMINARY; PRT; 278 AA.
AC Q8SQ34;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE CD40.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA West K.A., Li A.W., Rowden G.;
RT "Characterization of the Porcine CD40 Molecule.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF248545; AAL02924.1; ..
SQ SEQUENCE 278 AA; 30951 MW; 20D446E44AF93DD2 CRC64;
```

```
Query Match 13.5%; Score 303.5; DB 6; Length 278;
Best Local Similarity 35.8%; Pred. No. 1.5e-16;
Matches 72; Conservative 24; Mismatches 80; Indels 25; Gaps 8;

Qy 16 PLVLGLFG-LLAASQPAVPPVASENQCRCDOEYEEFQHRICSRCPGPGYVSAKCSR 74
Db 5 PLKLLWGCFLTAVHPE--PP-----TSCK--ENQY--PTNSRCCNLCPPGQKLVNHCTE 53
Qy 75 IRDVTCACTAENSNEHWNLYICOLCRPCDPMGLEETAPCTSKRKTCRCQCPGMFCAA 134
Db 54 VTETCLPCSSSEFLATWNREKHQHKYCDNLGLQVOREGTSKTDTCVCGSEGHHCYN 113
Qy 135 WALECTHCELLSDCPPG-----TEAELKDEVGKNNHCVPCKAGHFQNTSSPSARCQPH 189
Db 114 SA--CESCTLHSLCFFPLGLGVKQMADEVSDTI-----CEPCPVGFFSNVSSASERCPWT 165
Qy 190 RCENOGVLEAAPTQAQSDTTC 210
Db 166 SCESKGLVEQRAQTKNTDVC 186
```

```
RESULT 5
Q8WMQ2
ID Q8WMQ2 PRELIMINARY; PRT; 277 AA.
AC Q8WMQ2;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Membrane protein CD40 (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RA Zawitkowski M.S., Russ G.R., Krishnan R.;
```

```
RT "Cloning and expression of the ovine CD40 molecule and the inhibition
RT of the mixed lymphocyte reaction by the ovine CD40-EGFP fusion
RT protein.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY072798; AAL68402.1; ..
DR InterPro; IPR001368; TNFR_c6.
DR Pfam; PF00020; TNFR_c6; 4.
DR SMART; SM00208; TNFR_4.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS50050; TNFR_NGFR_2; 1.
FT NON_TER 277
SQ SEQUENCE 277 AA; 30465 MW; 619F28BEA18A0D29 CRC64;
```

```
Query Match 12.1%; Score 271; DB 6; Length 277;
Best Local Similarity 26.7%; Pred. No. 5.8e-14;
Matches 76; Conservative 29; Mismatches 100; Indels 80; Gaps 9;

Qy 47 EKEYEYEQHRICSRCPGPGYVSAKSRIRDTVCATCAENSNEHWNLYICOLCRPCPD 106
Db 28 EKQY--PVNSLCCDLCPGQKLLNDCTEYKTECQSGKGEFLSTWNRKYCHEHRYCNP 85
Qy 107 VMGLEEIAPCTSKRKTCRCQCPGMFCAAWALECTHCELLSDCPPGTEAELKDEVGK--N 164
Db 86 NLGLRIQSEGLTNTDTTCVDEGQHCTSH--CESCTPHSLCLCPGFGVK---QIATGVLD 140
Qy 165 NHCVPCKAGHFONTSSPSARCOPHTRCENOGVLEAAPTQAQSDTTC-----210
Db 141 TTCEPCPVGFFSNVSSAFKCHPWTSCKERKGLVGHQVGNKNDVAVCGFOSRMRTLVAIP 200
Qy 211 -----KNPLEPLPPMMSGSLKRR-----PQSGEPNPVAGSWEP 244
Db 201 TWGVLFVALLVSACIRNMTKKRQLRPCTLRKGRIPWRRILRIIPF---GPTRLSG--- 253
Qy 245 PRAHPFPDLVOPLLPISGDVSPVSTGLPAPVLEAGVQQQSPL 289
Db 254 -----AGDMLVSAAGP-----GGRQRESHL 274
```

```
RESULT 6
Q88SM6
ID Q88SM6 PRELIMINARY; PRT; 651 AA.
AC Q88SM6;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Death receptor 6.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Bridgman J.T., Johnson A.L.;
RT "Expression of DR6 in the ovary.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF349908; AAK29666.2; ..
DR HSP; P19438; INCF.
DR InterPro; IPR00488; Death.
DR InterPro; IPR003975; Shal_Channel.
DR InterPro; IPR001368; TNFR_c6.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00020; TNFR_c6; 4.
DR PRINTS; PR01497; SHALCHANNEL.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00208; TNFR; 4.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS50050; TNFR_NGFR_2; 1.
KW Receptor.
SQ SEQUENCE 651 AA; 71003 MW; BDC95A600DAB2C2A CRC64;
```

Query Match 11.8%; Score 266; DB 13; Length 651;

```
Best Local Similarity 25.1%; Pred. No. 3.9e-13;
Matches 107; Conservative 46; Mismatches 150; Indels 124; Gaps 25;

QY 18 VLGLFGLLAASQPAVPPYASENOTCRDQKEYEY-----PQHRICCSRCPPGTIVVSACKS 73
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9 VLPLLVFLGTADAQ--PKLTSEQANVSLPAGKYHLDRATNOELICDKCPAGTYVSKHCT 66
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 74 RIRDTVCATCAENSYNHWNHNYLTICQLCR-PCDPVMGLEETAPCTSKRKTQCRCQPGMF- 131
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 67 KSTLRKCSPCPDGTFTRKHGIERCHPCRPCE--LPWIEKTHCTALTRECTCLSGTFQ 124
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 132 ----CAWALECTHCELLSDCP-----PGTEAEKDEVGKGNHCVPCKAGHFQNTSS 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 125 INDRVCVPTV-----CPVGWYRKKGTETE-----DVRCKPCLRGTFSDVPS 166
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 181 PSARCQPHTRCENGLVEAAPTQSDTTCNPLPPPEMSGSLKRRRPPGEGPNPVAG 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 167 SVMCKTYTDFGKMMVVRPGTESDNCVGPAS-LP---NLSLSSDAQADETYEA- 221
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 241 SWEPKHAHYFP-DLVQPLPIGSDVSP-VSTGLPAPVLEAGVPQQOQSLDLTREPQLE 298
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 222 ---PPTA--YLPKGLNSVDFLSSSPAPRVSN-----TAEPD 256
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 299 PGESQVAGTNGTHVTGGSWTNGTNI-----YIYNGPVLGPPGP-----GDLPATPE 347
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 257 YNDTS---ANGTVG---APGSLSSAGTQAQSYRHKHTSQAMGKQPAQEMAGGEKSSIPY 311
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 348 PYPPIPEGDPGPPGLSTPHQEDKAWHLAETH-----CGATPSNR 389
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 312 RP---GPNV---HQ-----HFDINEHLPWMLVFLLLVLLVIVVCSVRSSR 356
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 390 ----GPR 392
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 357 TLKKGPR 363
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
Q9DDD2 ID Q9DDD2 PRELIMINARY; PRT; 276 AA.
AC Q9DDD2;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DE 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Human CD40-homologue.
GN TNFSF5.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Tregaskes C.A.;
RL Thesis (2001), University of Reading, Reading, UNITED KINGDOM.
DR EMBL; A2937700; CAC2018.1; -.
DR HSSP; P25942; ICDF.
DR InterPro; IPR001005; Myb_DNA_binding.
DR InterPro; IPR001368; TNFR_c6.
DR Pfam; PF00020; TNFR_c6; 3.
DR SMART; SM00208; TNFR; 3.
DR PROSITE; PS00037; MYB_1; UNKNOWN_1.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS00050; TNFR_NGFR_2; 3.
SQ SEQUENCE 276 AA; 30009 MW; 4040B7E0DB82454E CRC64;

Query Match 11.3%; Score 253; DB 13; Length 276;
Best Local Similarity 26.3%; Pred. No. 1.6e-12;
Matches 84; Conservative 31; Mismatches 122; Indels 82; Gaps 13;

QY 19 LGFLGLLA-----SOPQAVPPYASENOTCRDQKEYEYEQHRIICCSRCPPGTIVVSACKS 73
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4 LGLLGLLALLGGQGP-----GDAVNCSDQ---YEHKGR--CCNRCQPGKKLASECN 52
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
QY 74 RIRDTVCATCAENSYNHWNHNYLTICQLCRPCDPVMGLEETAPCTSKRKTQCRCQPGMFA 133
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 53 DTEDSVCTPCENGQYQHSWTKERHCTPHEICEDNAGLIVRHRGNATHNTVCQCRAGMHCS 112
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 134 AWALECTHCELLSDCPPG-----TEAEKDEVGKGNHCVPCKAGHFQNTSSPSARCO 186
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 113 --DASCOTCVENEPCKOGFGFVAAAEARM-----TSPCEPCAEGTFSNVSKTEPCH 163
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 187 PHTRCENQGLVEAAPTQSDTTCNKP-----LEPLPPPEMSGSL-----LK 227
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 164 FWISCEBGLVVKVKGINTSDVICESSRRSSLSVLIPITAAVVTCLVGICICVCLVHTDLR 223
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 228 RPOGEGPNVAGSWEPKHAHYFPDLVQLPLPIGSDVSPVSTGLPAPVLEAGVPQQOS 287
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 224 RR-----GPKQAEAEAPR-----ELVTQ-----QPEEVDFPVQETLLGGQPVQA- 262
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 288 PLDLTREPQLEPGEOSOVA 306
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 263 -----EDGESRIA 271
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
O02764 ID O02764 PRELIMINARY; PRT; 267 AA.
AC O02764;
DT 01-JUL-1997 (T-EMBLrel. 04, Created)
DT 01-JUL-1997 (T-EMBLrel. 04, Last sequence update)
DE 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE OX40 precursor (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Isono T., Seto A.;
RL "Expression of OX40 and OX40 ligand genes in rabbit HTLV-I-transformed T cell lines."
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB003911; BRA20059.1; -.
DR HSSP; P25942; ICDF.
DR InterPro; IPR001368; TNFR_c6.
DR Pfam; PF00020; TNFR_c6; 3.
DR SMART; SM00208; TNFR; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00050; TNFR_NGFR_2; 2.
KW Signal.
FT NON_TER 1 1
FT SIGNAL <1 18 POTENTIAL.
FT CHAIN 19 267 OX40.
SQ SEQUENCE 267 AA; 28489 MW; A8B4CD3173C9500B CRC64;

Query Match 11.2%; Score 252.5; DB 6; Length 267;
Best Local Similarity 26.7%; Pred. No. 1.7e-12;
Matches 75; Conservative 32; Mismatches 95; Indels 79; Gaps 11;

QY 17 LVGLFGLLAASQPAVPPYASENOTCRDQKEYEYEQHRIICCSRCPPGTIVVSACKSRI 76
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8 LALLLGLLIGAEP--PDCVGDTPGDR-----CCLECPQGYGMVSRCHRSQ 54
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 77 DTVCATCAENSYNHWNHNYLTICQLCRPCDPVMGLEETAPCTSKRKTQCRCQPGMFAAWA 136
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 55 DTICHPCPGFYNEAVNY-QACKPCTQCNRRSGSEPOQECTHTTRDTVCRCRPG----- 106
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 137 LECHCELLSDCPPGTEAEKDEVGKGNHCVPCKAGHFQNTSSPSARCPHTCENQGL 196
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 107 -----TQPLNGYKHGVD-----CAPCQGHFSEGNRA--CRPWTNCTLAGK 146
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 197 VEAPGTAQSDTTCNPLPPPEMSGSLKRRRPPGEGPNPVAGSWEPKHAHYFPDLVQ 256
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 147 RTLOPASSISDAVCE-----RSSLATQP-----WETPSA-PYRPTAR 184
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```



```

Db 70 TKERCPPHTYTPVNSYSGCHQCRKC-PTGSFDKV-KCTGTQNSKCSCLPGWFCATDSS 127
QY 135 WALECTHCELLSDCPGTEAEKLDKGVKGNHCVCKAGHF 175
Db 128 KYEDCRDCLPKRCKPCGCGYGGI--DEL--GNPLCKSCCVGEY 165

RESULT 14
O57117 PRELIMINARY; PRT; 351 AA.
AC O57117;
DT 01-JUN-1998 (TremBLrel. 06, Created)
DT 01-JUN-1998 (TremBLrel. 06, Last sequence update)
DE 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Tumor necrosis factor receptor II homolog.
GN CRMB.
OS Cowpox virus (CPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10243;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MUNICH OPV 85 (HUMAN);
RX MEDLINE=20379014; PubMed=11136755;
RA Loparev V.N., Massung R.F., Esposito J.J., Meyer H.;
RT "Detection and differentiation of old world orthopoxviruses:
RT restriction fragment length polymorphism of the crmb gene region.";
RL J. Clin. Microbiol. 39:94-100(2001).
DR EMBL; U90227; AB94383.1; -
DR HSP; O14763; IDOG.
DR InterPro; IPR001368; TNFR_c6.
DR Pfam; PF00020; TNFR_c6; 2.
DR SMART; SM00208; TNFR; 2.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_2.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
KW Receptor.
SQ SEQUENCE 351 AA; 38304 MW; 57C84FFE5B5F57E CRC64;

Query Match 10.2%; Score 230; DB 12; Length 351;
Best Local Similarity 31.4%; Pred. No. 1.5e-10;
Matches 50; Conservative 18; Mismatches 79; Indels 12; Gaps 6;

QY 33 VPPYASENCTDQKEYYEPQHRRICCSRCPPGTYVSARK--SRIRDTVCATCAENSYNE 90
Db 22 VAPYAPSNCKRCDNEYN----RHNLCLSCPPTGTYASRLCDSKTNTTQCTPGSGCTFTS 77
QY 91 HNNYLTICOLR-PCDPVVGLEETAPCTSKRKTQCRQPMFC-AAWALECTHCELLSDC 148
Db 78 HNNHLPACLSNCRCD--SNQVETRPCTNTHNRICBSPGYCYLLKGGSGCKACVSTK 135
QY 149 PCTEAEKLDKGVKGNHCVCKAGHFQNTSSPSARCP 187
Db 136 --GIGYVSGHTSTGDIVCSPCGLGTYSHTVSSADKCEP 172

RESULT 15
Q90YS6 PRELIMINARY; PRT; 285 AA.
AC Q90YS6;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE TNF decoy receptor.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RA Liu L., Fujiki K., Dixon B., Sundick R.S.;
RT "Cloning of a novel rainbow trout (Oncorhynchus mykiss) CC chemokine
```

```

RT with a fractalkine-like stalk and a TNF decoy receptor using cdna
RT fragments containing AU-rich elements.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF401631; AAK91758.1; -
DR InterPro; IPR000561; EGF-like.
DR Pfam; PF00020; TNFR_c6; 3.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS50050; TNFR_NGFR_2; 1.
KW Receptor.
SQ SEQUENCE 285 AA; 31795 MW; 5E3BD1B6EFC6BABC CRC64;

Query Match 10.1%; Score 228; DB 13; Length 285;
Best Local Similarity 29.6%; Pred. No. 1.6e-10;
Matches 61; Conservative 18; Mismatches 87; Indels 40; Gaps 10;

QY 17 LVGLGLFLAASQQA---VPPYASENCTDQKEYYEPQHRRICCSRCPPGTYVSARKS 73
Db 9 LPLLVFALCGGSPVPSGAHTPTVIWRDDATGDS-----LTCDLCAPGTYLLKHCT 58
QY 74 RIRDTVCATCAENSYNEHNNYLTICOLC-RPC--DPVVGLEETAPCTSKRKTQCRQPM 130
Db 59 KDRKSDCGPCPKSHYTEIWNYYIERCQYCNRFCTADEI----ESVPCTQLHNRQCECKDGF 114
QY 131 FCAAWALECTH--CELLSDCPPTGTEAEKLDKGVKGNH---CVPCKAGHFQNTSSPSAR 184
Db 115 Y-----MTHGSCSRHRCPPG-----EGVISNGTAHTDVKCEPCVGVFFSAVSSRKA 162
QY 185 CQPTRENOGLVEAAPGTAQSDTTC 210
Db 163 CQKFSVCPPGR--TTIPGNDMDNDVC 186

Search completed: April 8, 2003, 10:51:35
Job time : 105.845 secs
```


GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2003, 10:47:16 ; Search time 35.2339 Seconds
(without alignments)
333.195 Million cell updates/sec

Title: US-09-917-372-1
Perfect score: 2248
Sequence: 1 MLLPWATSAFGLAWPLVLG.....EHCATPSNRGPRNOFITHD 399

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2.6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2.6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2.6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2.6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2.6/ptodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2.6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1404	62.5	415	4	US-09-006-353A-6
2	1404	62.5	415	4	US-09-573-986-6
3	1129	50.2	197	2	US-08-505-606-1
4	1129	50.2	197	4	US-09-000-166-1
5	987	43.9	170	4	US-08-828-683A-14
6	456	20.3	77	4	US-08-866-545-3
7	349.5	15.5	461	1	US-08-385-223-2
8	349.5	15.5	461	2	US-08-650-000-2
9	349.5	15.5	461	4	US-09-042-785A-7
10	349.5	15.5	461	4	US-08-477-347-3
11	349.5	15.5	461	4	US-09-006-353A-4
12	349.5	15.5	461	4	US-08-476-862-2
13	349.5	15.5	461	4	US-09-573-986-4
14	349.5	15.5	461	6	5395760-2
15	339.5	15.1	518	1	US-08-385-229-4
16	337.5	15.0	486	1	US-08-243-010-1
17	327.5	14.6	474	2	US-08-650-000-4
18	327.5	14.6	474	4	US-09-042-785A-8
19	327.5	14.6	474	6	5395760-4
20	321.5	14.3	227	3	US-08-974-022-48
21	321.5	14.3	227	4	US-08-795-445A-48
22	321.5	14.3	227	4	US-08-795-447A-48
23	321.5	14.3	227	4	US-08-974-186-48
24	321.5	14.3	227	4	US-08-795-448B-48
25	321.5	14.3	227	4	US-08-706-945D-134
26	319	14.2	300	2	US-08-794-796-2
27	314.5	14.0	276	4	US-09-041-886-27
28	314.5	14.0	277	4	US-09-042-785A-10
29	314.5	14.0	277	4	US-09-006-353A-10
30	314.5	14.0	277	4	US-08-114-944D-2
31	314.5	14.0	277	4	US-09-573-986-10
32	312	13.9	625	3	US-08-996-139-15
33	312	13.9	625	4	US-08-995-659-15
34	312	13.9	625	4	US-09-215-649A-15
35	312	13.9	625	4	US-09-577-780-15
36	310.5	13.8	299	4	US-09-286-529-17
37	309.5	13.8	235	4	US-09-326-394-4
38	309.5	13.8	235	4	US-09-580-233-2
39	309.5	13.8	235	4	US-09-580-235-8
40	309.5	13.8	235	4	US-09-580-181-2
41	309.5	13.8	235	4	US-09-580-181-8
42	309.5	13.8	235	4	US-09-102-530-2
43	309.5	13.8	235	4	US-09-102-530-8
44	308.5	13.7	235	4	US-09-580-235-4
45	308.5	13.7	235	4	US-09-580-235-6

RESULT 1
US-09-006-353A-6
; Sequence 6, Application US/09006353A
; Patent No. 6261801
; GENERAL INFORMATION:
; APPLICANT: WEI, YING-FEI
; APPLICANT: YU, GUO-LIANG
; APPLICANT: GENTZ, REINER
; APPLICANT: RUBEN, STEVEN
; TITLE OF INVENTION: TUMOR NECROSIS FACTOR RECEPTOR 5
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/006,353A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF341
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 415 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-006-353A-6

Query Match 62.5%; Score 1404; DB 4; Length 415;

Best Local Similarity 63.9%; Pred. No. 4.7e-95;

Matches 265; Conservative 28; Mismatches 86; Indels 36; Gaps 4;

QY 1 MLLPWATSAFGLAWPLVLGFLGLAASQPAVPPYASENQTCRQKEKEYEPQHRICCS 60

DB 1 MRLPRASSPCGLAWGPLLLGLSGLLVASQPLVPPYRIENQTCWDQDKEYEYEPMDVCCS 60

Db 297 PAGPTAPSLVLEEVVLOQQSPVQARELEAPGEGHGVAGHANGIHVTGGSVTVTGNIIY 356

QY 329 NGVILGGPPGDLPATPEPPYPIPEGDPGPPGLSTPHOEDGKAWHLAETEHCG 383

Db 357 NGPVLGTRGPGDPAPPPEPPYPTPEGAPGPPSELSTPYQEDGKAWHLAETETLG 411

RESULT 3

US-08-505-606-1

; Sequence 1, Application US/08505606

; Patent No. 5925351

; GENERAL INFORMATION:

; APPLICANT: BROWNING, Jeffrey L.

; APPLICANT: BROWNING, Christopher D.

; APPLICANT: HOCJMAN, Paula S.

; TITLE OF INVENTION: SOLUBLE LYMPHOTOXIN-BETA RECEPTORS AND

; TITLE OF INVENTION: ANTI-LYMPHOTOXIN RECEPTOR AND LIGAND ANTIBODIES AS

; TITLE OF INVENTION: THERAPEUTIC AGENTS FOR THE TREATMENT OF IMMUNOLOGICAL

; NUMBER OF SEQUENCES: 1

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: James F. Haley, Jr.

; STREET: 1251 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10020

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/505,606

; FILING DATE: 21-JUL-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/378,968

; FILING DATE: 26-JAN-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: HALEY, Jr., James F.

; REGISTRATION NUMBER: 27,794

; REFERENCE/DOCKET NUMBER: B191

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 596-9000

; TELEFAX: (212) 596-9090

; TELEX: 14-8367

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 197 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-505-606-1

Query Match 50.2%; Score 1129; DB 2; Length 197;

Best Local Similarity 99.5%; Pred. No. 2.3e-75;

Matches 196; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 28 SOPQAVPPYASNQTRDOEKEYEYEPQHRICCSRCPPGTYVYSAKCSRIKRDIVCATCAENS 87

Db 1 SOPQAVPPYASNQTRDOEKEYEYEPQHRICCSRCPPGTYVYSAKCSRIKRDIVCATCAENS 60

QY 88 YNEHWNLYTICOLCRPCDPVVMGLEIAPCTSKRKTQCRQCPQCFCAAALECHTCELLSD 147

Db 61 YNEHWNLYTICOLCRPCDPVVMGLEIAPCTSKRKTQCRQCPQCFCAAALECHTCELLSD 120

QY 148 CPPGTAEALKDEVGKNNHCVCKAGHFQNTSSPARGQPHTRCENQGLVEAPGTAQSD 207

Db 121 CPPGTAEALKDEVGKNNHCVCKAGHFQNTSSPARGQPHTRCENQGLVEAPGTAQSD 180

QY 208 TTCKNPLEPLPPMSG 224

QY 61 RCPGTYVYSAKCSRIKRDIVCATCAENS YNEHWNLYTICOLCRPCDPVVMGLEIAPCTSKR 120

Db 61 RCPGTYVYSAKCSRIKRDIVCATCAENS YNEHWNLYTICOLCRPCDPVVMGLEIAPCTSKR 120

QY 121 KTCRCQPGMFCAAWALACTHC--ELLSDCPGTEAEELKDEVGKNNHCVCKAGHFQNT 178

Db 121 KTCRCQPGMFCAAWALACTHC--ELLSDCPGTEAEELKDEVGKNNHCVCKAGHFQNT 180

QY 179 SSPARCOPTRCENQGLVEAPGTAQSDTTCKNPLEP----- 216

Db 181 SSPARCOPTRCENQGLVEAPGTAQSDTTCKNPLEP----- 240

QY 217 -----LPPEMS-----GSLKRRPQEGNPNVAGSWEPKHAHYFPDLPVLPISGDSVPV 268

Db 241 ACAMWRHPSLCRKLGTLLKRRHPEGESPPCPA-----PRADPHFPDLAEPLPMSGDLSPS 296

QY 269 STGLPAAPVLEAGVQQQSPDLTREPQLEGEQSQVAHGTNGIHVTGSGMTITGNIIY 328

Db 297 PAGPTAPSLVLEEVVLOQQSPVQARELEAPGEGHGVAGHANGIHVTGGSVTVTGNIIY 356

QY 329 NGVILGGPPGDLPATPEPPYPIPEGDPGPPGLSTPHOEDGKAWHLAETEHCG 383

Db 357 NGPVLGTRGPGDPAPPPEPPYPTPEGAPGPPSELSTPYQEDGKAWHLAETETLG 411

RESULT 2

US-09-573-986-6

; Sequence 6, Application US/09573986

; Patent No. 6455040

; GENERAL INFORMATION:

; APPLICANT: Wei, Ying-Fei

; APPLICANT: Ni, Jian

; APPLICANT: Gentz, Reiner

; APPLICANT: Ruben, Steven

; TITLE OF INVENTION: Tumor Necrosis Factor Receptor 5

; FILE REFERENCE: 1488.1280004

; CURRENT APPLICATION NUMBER: US/09/573,986

; CURRENT FILING DATE: 2000-05-18

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-573-986-6

Query Match 62.5%; Score 1404; DB 4; Length 415;

Best Local Similarity 63.9%; Pred. No. 4.7e-95;

Matches 265; Conservative 28; Mismatches 86; Indels 36; Gaps 4;

QY 1 MLPLWATAPGLANGPLVLGLFLLAASOPQAVPPYASNQTRDOEKEYEYEPQHRICCS 60

Db 1 MLPLWATAPGLANGPLVLGLFLLAASOPQAVPPYASNQTRDOEKEYEYEPQHRICCS 60

QY 61 RCPGTYVYSAKCSRIKRDIVCATCAENS YNEHWNLYTICOLCRPCDPVVMGLEIAPCTSKR 120

Db 61 RCPGTYVYSAKCSRIKRDIVCATCAENS YNEHWNLYTICOLCRPCDPVVMGLEIAPCTSKR 120

QY 121 KTCRCQPGMFCAAWALACTHC--ELLSDCPGTEAEELKDEVGKNNHCVCKAGHFQNT 178

Db 121 KTCRCQPGMFCAAWALACTHC--ELLSDCPGTEAEELKDEVGKNNHCVCKAGHFQNT 180

QY 179 SSPARCOPTRCENQGLVEAPGTAQSDTTCKNPLEP----- 216

Db 181 SSPARCOPTRCENQGLVEAPGTAQSDTTCKNPLEP----- 240

QY 217 -----LPPEMS-----GSLKRRPQEGNPNVAGSWEPKHAHYFPDLPVLPISGDSVPV 268

Db 241 ACAMWRHPSLCRKLGTLLKRRHPEGESPPCPA-----PRADPHFPDLAEPLPMSGDLSPS 296

QY 269 STGLPAAPVLEAGVQQQSPDLTREPQLEGEQSQVAHGTNGIHVTGSGMTITGNIIY 328

	TELEFAX: 650-493-5556	
	TELEX: 66141 PENNIE	
	INFORMATION FOR SEQ ID NO: 3:	
	SEQUENCE CHARACTERISTICS:	
	LENGTH: 77 amino acids	
	TYPE: amino acid	
	STRANDEDNESS: single	
	TOPOLOGY: linear	
	MOLECULE TYPE: No. 6265535e	
	US-08-866-545-3	
	Query Match	20.3%; Score 456; DB 4; Length 77;
	Best Local Similarity	100.0%; Pred.No. 9.7e-27;
	Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	79 VCATCAENSYNEHWNITICQLCRPCDPVGMGLBEIAICTSKRKTQCRCQPGMFCFAAWALE 138 	
Db	1 VCATCAENSYNEHWNITICQLCRPCDPVGMGLBEIAICTSKRKTQCRCQPGMFCFAAWALE 60 	
QY	139 CTHCELLSDCPPTGEAE 155 	
Db	61 CTHCELLSDCPPTGEAE 77 	
	RESULT 7	
	US-08-385-229-2	
	; Sequence 2, Application US/08385229	
	; Patent No. 5605690	
	; GENERAL INFORMATION:	
	; APPLICANT: Jacobs, Cindy A.	
	; APPLICANT: Smith, Craig A.	
	; TITLE OF INVENTION: Method of Treating TNF-Dependent	
	; TITLE OF INVENTION: Inflammation Using Tumor Necrosis Factor Antagonists	
	; NUMBER OF SEQUENCES: 5	
	; CORRESPONDENCE ADDRESS:	
	; ADDRESSEE: Immunex Corporation	
	; STREET: 51 University Street	
	; CITY: Seattle	
	; STATE: Washington	
	; COUNTRY: U.S.A.	
	; ZIP: 98101	
	; COMPUTER READABLE FORM:	
	; MEDIUM TYPE: Floppy disk	
	; COMPUTER: IBM PC compatible	
	; OPERATING SYSTEM: PC-DOS/MS-DOS	
	; SOFTWARE: PatentIn Release #1.0, Version #1.25	
	; CURRENT APPLICATION DATA:	
	; APPLICATION NUMBER: US/08/385,229	
	; FILING DATE:	
	; CLASSIFICATION: 435	
	; PRIOR APPLICATION DATA:	
	; APPLICATION NUMBER: US/07/946,236	
	; FILING DATE:	
	; ATTORNEY/AGENT INFORMATION:	
	; NAME: Wight, Christopher L.	
	; REGISTRATION NUMBER: 31,680	
	; REFERENCE/DOCKET NUMBER: 2503	
	; TELECOMMUNICATION INFORMATION:	
	; TELEPHONE: (206) 587-0430	
	; TELEFAX: (206) 587-0606	
	; INFORMATION FOR SEQ ID NO: 2:	
	; SEQUENCE CHARACTERISTICS:	
	; LENGTH: 461 amino acids	
	; TYPE: amino acid	
	; TOPOLOGY: linear	
	; MOLECULE TYPE: protein	
	US-08-385-229-2	
	Query Match	15.5%; Score 349.5; DB 1; Length 461;
	Best Local Similarity	29.7%; Pred. No. 4.3e-18;
	Matches 113; Conservative 47; Mismatches 143; Indels 77; Gaps 22;	
Y	9 APGLANGPLVLGFLGLAASQFOAVP-----PYASE-NOTCRDQKEYYEPQHRIICCSK 61	

; REFERENCE/DOCKET NUMBER: 2501-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 461 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-650-000-2

Query Match 15.5%; Score 349.5; DB 2; Length 461;
Best Local Similarity 29.7%; Pred. No. 4.3e-18;
Matches 113; Conservative 47; Mismatches 143; Indels 77; Gaps 22;

QY 9 APGLAWGLVLGLFLLAASQPAVP-----PYASE-NQCRDQEKYYEPQHRCICSR 61
DB 2 APVAVMAALAVGL-ELWAAA--HALPAQVAFTPYAPEPGSTCR--LREYDQTAQMCCSK 56
QY 62 CPPTGYVSAKCSIRDTVCATCAENSYNEHNNYLTICOLCR---PCDPVMGLEETAPCTS 118
DB 57 CSPGQHAKVFCIKTSDTVCDSCEDSTYTLNWNVPECLSCGSRCSDDV---ETQACTR 112
QY 119 KRKTCRCQPMFCFAAWALE-CTHCELLSDCPPG-----TEAELKDEYKGNHNCVPCA 172
DB 113 EQNRICTRCPGWYKALSKQEGRCRLCAPLRCRPGFGVARPGTETSDVV-----CKPCAP 166
QY 173 GHFNTSSPSARCQPHTRCENGLVEAAPGTAQSDTTC--KNPLEPLPPMSSGSL----- 225
DB 167 GTFSNTTSTDICRPHQICN-----VVAIPGNASMDAVCTSTSPTRSMAP---GAVHLPQP 219
QY 226 LKRRPQEGPNVAGSWPPKHAH-----PYFPDLVQPLLPISGDVS-PVS--TGLPAAP 276
DB 220 VSTRSQHTQPT-----EPSTAPSTSFLLPMGPS--PPAEGSTGDFALPVGLIVGTALG 272
QY 277 VLEAGVPQ-----QQSPDLITREPQLE--PGEQSOVAHGNTGTHV-----TGGNMT 320
DB 273 LLIIGVNCVIMTVQKKKPLCLQREAKVPHLPADKARGTQGPQQHLLITAPSSSSSLE 332
QY 321 ITGNIYIYNGPVILGGPPGPG 340
DB 333 SSASALDRRAPTRNQPAQPG 352

RESULT 9

US-09-042-785A-7

; Sequence 7, Application US/09042785A
; Patent No. 6194151

GENERAL INFORMATION:

; APPLICANT: Busfield, Samantha J
; TITLE OF INVENTION: NOVEL MOLECULES OF THE TNF RECEPTOR SUPERFAMILY
; TITLE OF INVENTION: AND USES THEREFOR

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD, LLP

; STREET: 28 State Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; APPLICATION NUMBER: US/09/042.785A

; FILING DATE: 17-MAR-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/938,896

; FILING DATE: 26-SEP-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Mandragouras, Amy E
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: MEI-001CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400

; TELEFAX: (617)742-4214

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 461 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; FRAGMENT TYPE: internal

US-09-042-785A-7

Query Match 15.5%; Score 349.5; DB 4; Length 461;
Best Local Similarity 29.7%; Pred. No. 4.3e-18;
Matches 113; Conservative 47; Mismatches 143; Indels 77; Gaps 22;

QY 9 APGLAWGLVLGLFLLAASQPAVP-----PYASE-NQCRDQEKYYEPQHRCICSR 61
DB 2 APVAVMAALAVGL-ELWAAA--HALPAQVAFTPYAPEPGSTCR--LREYDQTAQMCCSK 56
QY 62 CPPTGYVSAKCSIRDTVCATCAENSYNEHNNYLTICOLCR---PCDPVMGLEETAPCTS 118
DB 57 CSPGQHAKVFCIKTSDTVCDSCEDSTYTLNWNVPECLSCGSRCSDDV---ETQACTR 112
QY 119 KRKTCRCQPMFCFAAWALE-CTHCELLSDCPPG-----TEAELKDEYKGNHNCVPCA 172
DB 113 EQNRICTRCPGWYKALSKQEGRCRLCAPLRCRPGFGVARPGTETSDVV-----CKPCAP 166
QY 173 GHFNTSSPSARCQPHTRCENGLVEAAPGTAQSDTTC--KNPLEPLPPMSSGSL----- 225
DB 167 GTFSNTTSTDICRPHQICN-----VVAIPGNASMDAVCTSTSPTRSMAP---GAVHLPQP 219
QY 226 LKRRPQEGPNVAGSWPPKHAH-----PYFPDLVQPLLPISGDVS-PVS--TGLPAAP 276
DB 220 VSTRSQHTQPT-----EPSTAPSTSFLLPMGPS--PPAEGSTGDFALPVGLIVGTALG 272
QY 277 VLEAGVPQ-----QQSPDLITREPQLE--PGEQSOVAHGNTGTHV-----TGGNMT 320
DB 273 LLIIGVNCVIMTVQKKKPLCLQREAKVPHLPADKARGTQGPQQHLLITAPSSSSSLE 332
QY 321 ITGNIYIYNGPVILGGPPGPG 340
DB 333 SSASALDRRAPTRNQPAQPG 352

RESULT 10

US-08-477-347-3

; Sequence 3, Application US/08477347

; Patent No. 6232446

GENERAL INFORMATION:

; APPLICANT: WALLACH, David

; APPLICANT: BIGDA, Jacek

; APPLICANT: BELETSKY, Igor

; APPLICANT: METT, Igor

; TITLE OF INVENTION: TNF LIGANDS

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BROWDY AND NEIMARK

; STREET: 419 Seventh Street, N.W.

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20004

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/477,347

ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,862
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 107267
FILING DATE: 12-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 94039
FILING DATE: 06-APR-1990
APPLICATION NUMBER: IL 91229
FILING DATE: 06-AUG-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 90339
FILING DATE: 18-MAY-1989
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: WALLACH-12A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 461 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-476-862-2

Query Match 15.5%; Score 349.5; DB 4; Length 461;
Best Local Similarity 29.7%; Pred. No. 4.3e-18;
Matches 113; Conservative 47; Mismatches 143; Indels 77; Gaps 22;

QY 9 APGLAWGLVLGLFGLLAASQPAVP-----PYASE-NQTCROEKEYEYEPQHRICCSR 61
DB 2 APVAVWAALAVGL-ELWAAA--HALPAQVAFTPYAPEPGSTCR--LREYDQTAQMCCSK 56
QY 62 CPPGTIVYSAKCSRIKDTVCATCAENSNEHWNLTICOLCR---PCDPVVMGLEETAPCTS 118
DB 57 CSPQGHAKVFTKTSIDVDCSCESTYTLWNWPECLSCGSRSSDQV----ETQACTR 112
QY 119 KRKTQCRQCPGMFCAAWALE-CTHCELLSDCPPG-----TEAEKDEYKGNHNCVCPKA 172
DB 113 EONRICTCRPGWYKALSKEGCRCLCAPLRCRPGFVGARPGTETSDVV-----CKPCAP 166
QY 173 GHFONTSPPSARCOPHRCENOGIVEAPGTAQSDTTC--KNPLEPLPPMSGSL----- 225
DB 167 GTFSNTTSSDTCRPHQICN---VVAIPGNASMDAVCTSTSPTRSMAP---GAVHLPQP 219
QY 226 LKRRPQEGPNPVAGSWEPPKAH-----PYFPDLVQPLPISGDVS-PVS--TGLPAAP 276
DB 220 VSTRSQHTQPTP-----EPSTAPSTSLPMPGPS--PPAEGSTGDFALPVLGVGTALG 272
QY 277 VLEAGVPO-----QQSPDLTREPQLE--PGEQSQVAHGTNGIHV-----TGGSM 320
DB 273 LLIIGVNCVIMTQVKKKPLCLQREAKVPHLPADKARGTQGEQHQHLLITAPSSSSSLE 332
QY 321 ITGNIYYNGVPLGGPPGP 340
DB 333 SSASALDRRAPTRNQAPG 352

RESULT 13
US-09-573-986-4
Sequence 4, Application US/09573986
Patent No. 6455040
GENERAL INFORMATION:
APPLICANT: Wei, Ying-Fei
APPLICANT: Ni, Jian
APPLICANT: Gentz, Reiner
APPLICANT: Ruben, Steven
TITLE OF INVENTION: Tumor Necrosis Factor Receptor 5
FILE REFERENCE: 1488.1280004
CURRENT APPLICATION NUMBER: US/09/573,986
CURRENT FILING DATE: 2000-05-18
NUMBER OF SEQ ID NOS: 27
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 461
TYPE: PRT
ORGANISM: Homo sapiens
US-09-573-986-4

Query Match 15.5%; Score 349.5; DB 4; Length 461;
Best Local Similarity 29.7%; Pred. No. 4.3e-18;
Matches 113; Conservative 47; Mismatches 143; Indels 77; Gaps 22;

QY 9 APGLAWGLVLGLFGLLAASQPAVP-----PYASE-NQTCROEKEYEYEPQHRICCSR 61
DB 2 APVAVWAALAVGL-ELWAAA--HALPAQVAFTPYAPEPGSTCR--LREYDQTAQMCCSK 56
QY 62 CPPGTIVYSAKCSRIKDTVCATCAENSNEHWNLTICOLCR---PCDPVVMGLEETAPCTS 118
DB 57 CSPQGHAKVFTKTSIDVDCSCESTYTLWNWPECLSCGSRSSDQV----ETQACTR 112
QY 119 KRKTQCRQCPGMFCAAWALE-CTHCELLSDCPPG-----TEAEKDEYKGNHNCVCPKA 172
DB 113 EONRICTCRPGWYKALSKEGCRCLCAPLRCRPGFVGARPGTETSDVV-----CKPCAP 166
QY 173 GHFONTSPPSARCOPHRCENOGIVEAPGTAQSDTTC--KNPLEPLPPMSGSL----- 225
DB 167 GTFSNTTSSDTCRPHQICN---VVAIPGNASMDAVCTSTSPTRSMAP---GAVHLPQP 219
QY 226 LKRRPQEGPNPVAGSWEPPKAH-----PYFPDLVQPLPISGDVS-PVS--TGLPAAP 276
DB 220 VSTRSQHTQPTP-----EPSTAPSTSLPMPGPS--PPAEGSTGDFALPVLGVGTALG 272
QY 277 VLEAGVPO-----QQSPDLTREPQLE--PGEQSQVAHGTNGIHV-----TGGSM 320
DB 273 LLIIGVNCVIMTQVKKKPLCLQREAKVPHLPADKARGTQGEQHQHLLITAPSSSSSLE 332
QY 321 ITGNIYYNGVPLGGPPGP 340
DB 333 SSASALDRRAPTRNQAPG 352

RESULT 14
US-09-573-986-2
Patent No. 5395760
APPLICANT: SMITH, CRAIG A.; GOODWIN, RAYMOND G.; BECKMANN, M. PATRICIA
TITLE OF INVENTION: DNA ENCODING TUMOR NECROSIS FACTOR-a AND B-RECEPTORS
NUMBER OF SEQUENCES: 17
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/523,635
FILING DATE: 10-MAY-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 421,417
FILING DATE: 13-OCT-1989
APPLICATION NUMBER: 405,370
FILING DATE: 11-SEP-1989
APPLICATION NUMBER: 403,241

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2003, 10:45:16 ; Search time 5.15513 seconds
(without alignments)
799.386 Million cell updates/sec

Title: US-09-917-372-1_COPY_216_235
Perfect score: 108
Sequence: 1 PLPPMSGSLKRRPQGGP 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	46.3	125	17 Q9YCW4	Q9YCW4 aeropyrum p
2	49	45.4	139	11 Q9CTN9	Q9CTN9 mus musculus
3	49	45.4	688	5 Q9W111	Q9W111 drosophila
4	49	45.4	891	5 Q9VGJ8	Q9VGJ8 drosophila
5	49	45.4	1253	5 Q8WQ61	Q8WQ61 drosophila
6	49	45.4	2157	4 Q95875	Q95875 homo sapien
7	49	45.4	2157	4 Q960C6	Q960C6 homo sapien
8	48	44.4	930	11 Q925G9	Q925G9 mus musculus
9	47.5	44.0	1200	10 Q9ARZ5	Q9ARZ5 linum usita
10	47	43.5	618	10 Q9MA81	Q9MA81 arabidopsis
11	47	43.5	820	10 Q8RXZ0	Q8RXZ0 arabidopsis
12	47	43.5	1055	10 Q9MA51	Q9MA51 arabidopsis
13	47	43.5	1542	5 Q9VE07	Q9VE07 drosophila
14	46.5	43.1	511	6 Q95J00	Q95J00 sus scrofa
15	46.5	43.1	566	6 Q95JD1	Q95JD1 sus scrofa
16	46.5	43.1	572	16 Q33009	Q33009 mycobacteri

17	46.5	43.1	676	6 Q95JC9	Q95JC9 sus scrofa
18	46	42.6	471	10 Q49326	Q49326 arabidopsis
19	46	42.6	501	16 Q55762	Q55762 synecocyst
20	46	42.6	512	16 Q8XQ09	Q8XQ09 raietonia s
21	46	42.6	528	10 Q92PL6	Q92PL6 nicotiana t
22	46	42.6	598	4 Q8TBS5	Q8TBS5 homo sapien
23	46	42.6	645	10 Q94LV8	Q94LV8 oryza sativ
24	46	42.6	667	5 Q9Y174	Q9Y174 trichomonas
25	46	42.6	820	4 Q00236	Q00236 homo sapien
26	46	42.6	965	10 Q8S2E8	Q8S2E8 oryza sativ
27	46	42.6	1124	4 Q9Y6S2	Q9Y6S2 homo sapien
28	46	42.6	1124	4 Q99699	Q99699 homo sapien
29	46	42.6	1543	11 Q8R4H2	Q8R4H2 mus musculu
30	46	42.6	1544	4 Q9NZN5	Q9NZN5 homo sapien
31	45.5	42.1	207	13 Q9NL04	Q9NL04 eptatretus
32	45.5	42.1	363	11 Q62314	Q62314 mus musculu
33	45.5	42.1	754	11 Q9JJJ4	Q9JJJ4 mus musculu
34	45.5	42.1	754	11 Q99MU9	Q99MU9 mus musculu
35	45.5	42.1	754	11 Q9Z0R0	Q9Z0R0 mus musculu
36	45.5	42.1	1042	5 Q9V8C2	Q9V8C2 drosophila
37	45.5	42.1	1043	5 Q8T8P9	Q8T8P9 drosophila
38	45.5	42.1	1963	6 Q28019	Q28019 bos taurus
39	45	41.7	132	16 Q8TP91	Q8TP91 anabaena sp
40	45	41.7	210	16 Q9X9Z1	Q9X9Z1 streptomyce
41	45	41.7	233	4 Q9N0E2	Q9N0E2 homo sapien
42	45	41.7	337	10 Q8S0P8	Q8S0P8 oryza sativ
43	45	41.7	346	4 Q9NWY4	Q9NWY4 homo sapien
44	45	41.7	359	10 Q9AV86	Q9AV86 oryza sativ
45	45	41.7	397	16 Q9KY75	Q9KY75 streptomyce

ALIGNMENTS

RESULT 1
Q9YCW4 ID Q9YCW4 PRELIMINARY; PRT; 125 AA.
AC Q9YCW4;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein APE1148.
GN APE1148.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=9310339; PubMed=10382966;
RA Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anka A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000060; BAA80133.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 125 AA; 13011 MW; 5EA3CF9F929F1B21 CRC64;

Query Match 46.3%; Score 50; DB 17; Length 125;
Best Local Similarity 54.5%; Pred. No. 5.2;
Matches 12; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

QY 1 PLPPMSG--SLKRRPQGGP 20
|||:| |||:| |||:| |||:|
Db 10 PLLPQDDGRPLLRPGGGP 31

[illegible]

RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Beriman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fosler K., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RA "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY.
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacle J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celiniker S.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003692; AAF54681.1; -
 DR EMBL; AY070944; AAL48566.1; -
 DR FlyBase; FBgn0037934; CG6830.
 DR InterPro; IPR004119; DUF227.
 DR Pfam; PF02958; DUF227.2;
 SQ SEQUENCE 891 AA; 101884 MW; B723623AB4561563 CRC64;

 Query Match 45.4%; Score 49; DB 5; Length 891;
 Best Local Similarity 50.0%; Pred. No. 48;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

 QY 1 LPPPMGSLKRRPQGE 18
 ||||| :| :|||
 Db 460 PLPELLAALKPQGPQE 477

 RESULT 5
 Q8WQ61 PRELIMINARY; PRT; 1253 AA.
 ID Q8WQ61
 AC Q8WQ61
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Eps-15 protein.
 GN Eps-15.
 OS *Drosophila melanogaster* (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Walikar Y., Robinson I.M., O'Kane C.J.;
 RT "Characterisation of *Drosophila* Eps15 gene."
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ421624; CAD13484.1; -
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR000261; EPS15_repeat.
 DR Pfam; PF00036; efhand; 4.
 DR SMART; SM00054; EFh; 4.
 DR SMART; SM00027; EH; 3.
 DR PROSITE; PS00018; EF_HAND; UNKNOWN_2.
 SQ SEQUENCE 1253 AA; 134440 MW; 54CC8C62BEF8FECEB CRC64;

 Query Match 45.4%; Score 49; DB 5; Length 1253;
 Best Local Similarity 44.4%; Pred. No. 67;
 Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

 QY 2 LPPPMGSLKRRPQGE 19
 ||| :| :| :|||
 Db 228 MPPPPAGAMPAPSGEG 245

 RESULT 6
 O95875 PRELIMINARY; PRT; 2157 AA.
 ID O95875
 AC O95875;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE BAT2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Rowen L., Madan A., Qin S., Shaffer T., James R., Ratcliffe A.,
 RA Abbasi N., Dickhoff R., Loretz C., Madan A., Dors M., Young J.,
 RA Lasky S., Hood L.;
 RT "Sequence of the human major histocompatibility complex class III
 region."
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF129756; AAD18086.1; -
 DR EMBL; AF129756; AAD18086.1; -
 SQ SEQUENCE 2157 AA; 228869 MW; 21B817F5B699B0DB CRC64;

 Query Match 45.4%; Score 49; DB 4; Length 2157;
 Best Local Similarity 47.4%; Pred. No. 11e+02;
 Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

 QY 2 LPPPMGSLKRRPQGE 20
 ||| | :| :|||
 Db 1576 LPPHSSGFLGSKPEGPGP 1594

 RESULT 7
 Q96QC6 PRELIMINARY; PRT; 2157 AA.
 ID Q96QC6
 AC Q96QC6
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE BAT2 protein.
 GN BAT2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

OC Sequences from N.A.
 OC HIRAKAWA M., Yamaguchi H., Imai K., Shimada J.;
 OC Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN NCBI_TaxID=4006;
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, AKMOLINSK;
 RX MEDLINE=21096393; PubMed=11158537;
 RA DODDS P.N., Lawrence G.J., Ellis J.G.;
 RT "Six Amino Acid Changes Confined to the Leucine-Rich Repeat beta-
 RL Strand/Beta-Turn Motif Determine the Difference between the P and P2
 RT Rust Resistance Specificities in Flax.";
 RL Plant Cell 13:163-178 (2001).
 DR EMBL: AF310958; AAK28803.1;
 DR InterPro: IPR000767; Disease_resist.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR002182; NB-ARC.
 DR InterPro: IPR000157; TIR_domain.
 DR Pfam: PF00560; LRR; 8.
 DR Pfam: PF00931; NB-ARC; 1.
 DR Pfam: PF01582; TIR; 1.
 DR PRINTS: PR00364; DISEASERSIST.
 DR SMART: SM00255; TIR; 1.
 SQ SEQUENCE 1200 AA; 136588 MW; 09739254B3117E05 CRC64;

OC Sequences from N.A.
 OC HIRAKAWA M., Yamaguchi H., Imai K., Shimada J.;
 OC Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN NCBI_TaxID=4006;
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, AKMOLINSK;
 RX MEDLINE=21096393; PubMed=11158537;
 RA DODDS P.N., Lawrence G.J., Ellis J.G.;
 RT "Six Amino Acid Changes Confined to the Leucine-Rich Repeat beta-
 RL Strand/Beta-Turn Motif Determine the Difference between the P and P2
 RT Rust Resistance Specificities in Flax.";
 RL Plant Cell 13:163-178 (2001).
 DR EMBL: AF310958; AAK28803.1;
 DR InterPro: IPR000767; Disease_resist.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR002182; NB-ARC.
 DR InterPro: IPR000157; TIR_domain.
 DR Pfam: PF00560; LRR; 8.
 DR Pfam: PF00931; NB-ARC; 1.
 DR Pfam: PF01582; TIR; 1.
 DR PRINTS: PR00364; DISEASERSIST.
 DR SMART: SM00255; TIR; 1.
 SQ SEQUENCE 1200 AA; 136588 MW; 09739254B3117E05 CRC64;

Query Match 44.0%; Score 47.5; DB 10; Length 1200;
 Best Local Similarity 55.0%; Pred. No. 1.1e-02;
 Matches 11; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

Query Match 45.4%; Score 49; DB 4; Length 2157;
 Best Local Similarity 47.4%; Pred. No. 1.1e-02;
 Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 LPPEMSGSLKRRPQEGP 19
 |||||: ||| ||| |||
 DB 741 PLPPKLDLSKLLKIVRMQGLG 760

QY 2 LPPEMSGSLKRRPQEGP 20
 ||| | | : ||| |||
 DB 1576 LPPEMSGSLKRRPQEGP 1594

RESULT 8
 Q925G9 PRELIMINARY; PRT; 930 AA.
 AC Q925G9;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE PDZ-RGS3 protein.
 GN RGS3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21197945; PubMed=11301003;
 RA LU Q., Sun E.E., Klein R.S., Flanagan J.G.;
 RT "Phrin-b reverse signaling is mediated by a novel PDZ-RGS protein and
 RT selectively inhibits G protein-coupled chemoattraction.";
 RL Cell 105:69-79 (2001).
 DR EMBL: AF350047; AAK38878.1;
 DR MGD: MGI:1354734; Rgs3.
 DR InterPro: IPR001478; PDZ.
 DR InterPro: IPR001849; PH.
 DR InterPro: IPR000342; Regl_Gprotein.
 DR Pfam: PF00595; PDZ; 1.
 DR Pfam: PF00615; RGS; 1.
 DR PRODOM: PD001580; Reg_of_prg; 1.
 DR SMART: SM00233; PH; 1.
 DR PROSITE: PS0106; PDZ; 1.
 DR PROSITE: PS0132; RGS; 1.
 SQ SEQUENCE 930 AA; 102511 MW; FBOECEAL164D2F6A1 CRC64;

Query Match 44.4%; Score 48; DB 11; Length 930;
 Best Local Similarity 50.0%; Pred. No. 71;
 Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 LPPEMSGSLKRRPQEGP 20
 | : ||| | : ||| |||
 DB 535 PVPESSASVQKRLPSQESP 554

Query Match 44.4%; Score 47; DB 10; Length 618;
 Best Local Similarity 60.0%; Pred. No. 69;
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 LPPEMSGSLKRRPQ 16
 | : ||| ||| |||
 DB 476 LQPTSGSLKRRPK 490

Query Match 43.5%; Score 47; DB 10; Length 618;
 Best Local Similarity 60.0%; Pred. No. 69;
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

RESULT 9
 Q9ATZ5 PRELIMINARY; PRT; 1200 AA.
 AC Q9ATZ5;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Resistance-like protein P1-A.
 GN P1-A.
 OS Linum usitatissimum (Flax) (Linseed).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Query Match 44.4%; Score 48; DB 11; Length 930;
 Best Local Similarity 50.0%; Pred. No. 71;
 Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 LPPEMSGSLKRRPQEGP 20
 | : ||| | : ||| |||
 DB 535 PVPESSASVQKRLPSQESP 554

Query Match 44.4%; Score 48; DB 11; Length 930;
 Best Local Similarity 50.0%; Pred. No. 71;
 Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 LPPEMSGSLKRRPQEGP 19
 |||||: ||| ||| |||
 DB 741 PLPPKLDLSKLLKIVRMQGLG 760

Query Match 44.4%; Score 48; DB 11; Length 930;
 Best Local Similarity 50.0%; Pred. No. 71;
 Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 LPPEMSGSLKRRPQ 16
 | : ||| ||| |||
 DB 476 LQPTSGSLKRRPK 490

Query Match 43.5%; Score 47; DB 10; Length 618;
 Best Local Similarity 60.0%; Pred. No. 69;
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

RESULT 10
 Q9MA81 PRELIMINARY; PRT; 618 AA.
 AC Q9MA81;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE T12H1.35 protein (Fragment).
 GN T12H1.35.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=CV, COLUMBIA;
 RA LIN X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
 RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
 RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
 RT "Arabidopsis thaliana chromosome III BAC T12H1 genomic sequence.";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -!- SIMILARITY: CONTAINS 1 MYB-LIKE DOMAIN.
 DR EMBL: AC009177; AAF27044.1;
 DR InterPro: IPR001005; Myb_DNA_binding.
 DR Pfam: PF00249; myb_DNA_binding; 1.
 DR SMART: SM00395; SANT; 1.
 DR DNA-binding; Nuclear protein.
 KW NON_TER 618
 FT NON_TER 618
 SQ SEQUENCE 618 AA; 68924 MW; 234EE798627BA164 CRC64;

Query Match 43.5%; Score 47; DB 10; Length 618;
 Best Local Similarity 60.0%; Pred. No. 69;
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 LPPEMSGSLKRRPQ 16
 | : ||| ||| |||
 DB 476 LQPTSGSLKRRPK 490

Query Match 43.5%; Score 47; DB 10; Length 618;
 Best Local Similarity 60.0%; Pred. No. 69;
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```
RESULT 11
Q8RXZ0
ID Q8RXZ0 PRELIMINARY; PRT; 820 AA.
AC Q8RXZ0
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 91.8 kDa protein (Fragment).
GN AT3G05380
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kaniya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinzaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Arabidopsis Full Length cDNA Clones.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY080604; AAL86288.1; -
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 820 AA; 91757 MW; 98267CA802334008 CRC64;

Query Match 43.5%; Score 47; DB 10; Length 820;
Best Local Similarity 60.0%; Pred. No. 90;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 LPPEMSGSLRRPQ 16
I | : ||||| :| :
Db 241 LQPQTSGSLRRKPK 255

RESULT 12
Q9MA51
ID Q9MA51 PRELIMINARY; PRT; 1055 AA.
AC Q9MA51;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE F22F7.18 protein.
GN F22F7.18.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Uterback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana Chromosome III BAC F22F7 genomic sequence.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 MYB-LIKE DOMAIN.
DR EMBL; AC009606; AAF64544.1; -
DR InterPro; IPR001005; MYB_DNA_binding.
DR Pfam; PF00249; myb_DNA-binding; 1.
DR SMART; SM00395; SANT; 1.
DR SMART; SM00333; TUDOR; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 1055 AA; 118243 MW; F7B3A455BF5B1475 CRC64;
```

```
Query Match 43.5%; Score 47; DB 10; Length 1055;
Best Local Similarity 60.0%; Pred. No. 1.1e+02;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 LPPEMSGSLRRPQ 16
I | : ||||| :| :
Db 476 LQPQTSGSLRRKPK 490

RESULT 13
Q9VE07
ID Q9VE07 PRELIMINARY; PRT; 1542 AA.
AC Q9VE07;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CG6026 protein.
GN CG6026.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=2019606; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattie B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclev J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Testor C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003724; AAF55629.2; -
DR FlyBase; FBgn0038676; CG6026.
DR InterPro; IPR002965; P-rich_extensn.
DR PRINTS; PR01217; PRICHTEXTENS
SQ SEQUENCE 1542 AA; 168644 MW; 27C06E61A73FF016 CRC64;

Query Match 43.5%; Score 47; DB 5; Length 1542;
Best Local Similarity 52.9%; Pred. No. 1.6e+02;
```


GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM.protein - protein search, using sw model

Run on: April 8, 2003, 10:47:16 ; Search time 1.76611 seconds
(without alignments)
333.195 Million cell updates/sec

Title: US-09-917-372-1_COPY_216_235
Perfect score: 108
Sequence: 1 PLPPMSGSLKRRPOGEGP 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCRUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	42.6	197	US-08-505-606-1	Sequence 1, Appl
2	46	42.6	137	US-09-000-166-1	Sequence 1, Appl
3	46	42.6	1064	US-08-357-598-2	Sequence 2, Appl
4	46	42.6	1064	US-09-003-289-2	Sequence 2, Appl
5	46	42.6	1064	PCT-US95-16435-2	Sequence 2, Appl
6	46	42.6	1082	US-08-357-598-5	Sequence 5, Appl
7	46	42.6	1082	US-08-357-598-10	Sequence 10, Appl
8	46	42.6	1082	US-09-003-289-5	Sequence 5, Appl
9	46	42.6	1082	US-09-003-289-10	Sequence 10, Appl
10	46	42.6	1082	PCT-US95-16435-5	Sequence 5, Appl
11	46	42.6	1082	PCT-US95-16435-10	Sequence 10, Appl
12	46	42.6	1124	US-09-191-786-1	Sequence 1, Appl
13	46	42.6	1461	US-08-993-228-10	Sequence 10, Appl
14	45	41.7	415	US-09-006-353A-6	Sequence 6, Appl
15	45	41.7	415	US-09-573-986-6	Sequence 6, Appl
16	45	41.7	477	US-09-480-921B-10	Sequence 10, Appl
17	44	40.7	477	US-09-449-335-2	Sequence 2, Appl
18	44	40.7	477	US-09-449-335-6	Sequence 6, Appl
19	44	40.7	477	US-09-480-921B-29	Sequence 29, Appl
20	42	38.9	185	US-08-950-720A-13	Sequence 13, Appl
21	42	38.9	193	US-08-106-507-2	Sequence 2, Appl
22	42	38.9	302	US-08-893-853-3	Sequence 3, Appl
23	42	38.9	302	US-09-113-921-3	Sequence 3, Appl
24	42	38.9	894	US-08-117-362-4	Sequence 4, Appl
25	42	38.9	894	US-08-486-924-4	Sequence 4, Appl
26	42	38.9	1098	US-08-946-994-17	Sequence 17, Appl
27	42	38.9	1099	US-08-665-574C-16	Sequence 16, Appl

28	42	38.9	1100	1	US-08-357-598-11	Sequence 11, Appl
29	42	38.9	1100	2	US-09-003-289-11	Sequence 11, Appl
30	42	38.9	1100	5	PCT-US95-16435-11	Sequence 11, Appl
31	42	38.9	1299	5	PCT-US95-08354A-2	Sequence 2, Appl
32	42	38.9	1958	1	US-07-945-283-2	Sequence 2, Appl
33	41	38.0	563	4	US-09-422-936-79	Sequence 79, Appl
34	41	38.0	716	2	US-08-766-982-1	Sequence 1, Appl
35	41	38.0	716	4	US-09-296-219-1	Sequence 1, Appl
36	41	38.0	844	4	US-09-422-936-51	Sequence 51, Appl
37	41	38.0	886	4	US-09-422-936-77	Sequence 77, Appl
38	41	38.0	892	4	US-09-422-936-75	Sequence 75, Appl
39	41	38.0	899	4	US-09-422-936-71	Sequence 71, Appl
40	41	38.0	961	4	US-09-422-936-49	Sequence 49, Appl
41	40.5	37.5	1036	4	US-08-891-640-3	Sequence 3, Appl
42	40.5	37.5	1061	4	US-08-701-154A-5	Sequence 5, Appl
43	40.5	37.5	1336	2	US-08-231-193A-58	Sequence 58, Appl
44	40.5	37.5	1336	2	US-08-486-273A-58	Sequence 58, Appl
45	40.5	37.5	1336	3	US-08-940-086A-58	Sequence 58, Appl

ALIGNMENTS

RESULT 1
US-08-505-606-1
; Sequence 1, Application US/08505606
; Patent No. 5925351
; GENERAL INFORMATION:
; APPLICANT: BROWNING, Jeffrey L.
; APPLICANT: BENJAMIN, Christopher D.
; APPLICANT: HOCHMAN, Paula S.
; TITLE OF INVENTION: SOLUBLE LYMPHOTOXIN-BETA RECEPTORS AND
; TITLE OF INVENTION: ANTI-LYMPHOTOXIN RECEPTOR AND LIGAND ANTIBODIES AS
; TITLE OF INVENTION: THERAPEUTIC AGENTS FOR THE TREATMENT OF IMMUNOLOGICAL
; TITLE OF INVENTION: DISEASE
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr.
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/505,606
; FILING DATE: 21-JUL-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/378,968
; FILING DATE: 26-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY, Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B191
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 197 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-505-606-1

Query Match 42.6%; Score 46; DB 2; Length 197;

Db 419 PLGPDYKGLIRRSPTG 435

RESULT 11

PCT-US95-16435-10

Sequence 10, Application PC/TUS9516435

GENERAL INFORMATION:

APPLICANT: The Johns Hopkins University School of Medicine

TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 4225 Executive Square, Suite 1400

CITY: La Jolla

STATE: CA

COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/16435

FILING DATE: 15-DEC-1995

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Haile, Lisa A.

REGISTRATION NUMBER: 38,347

REFERENCE/DOCKET NUMBER: 07265/033W01

TELEPHONE: 619/678-5070

TELEFAX: 619/678-5099

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 1082 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US95-16435-10

Query Match 42.6%; Score 46; DB 5; Length 1082;

Best Local Similarity 47.1%; Pred. No. 72;

Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 PLPPMSGSLKRRPQG 17

|||:|:|:|

Db 419 PLGPDYKGLIRRSPTG 435

RESULT 12

US-09-191-786-1

Sequence 1, Application US/09191786

Patent No. 6372898

GENERAL INFORMATION:

APPLICANT: Cacalano, Nicholas A.

APPLICANT: Johnston, James A.

TITLE OF INVENTION: Mammalian Protein Variants and Methods

TITLE OF INVENTION: of Use

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: DNAX Research Institute

STREET: 901 California Avenue

CITY: Palo Alto

STATE: California

COUNTRY: USA

ZIP: 94304-1104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/993,228

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Sprunger, Suzanne A.

REGISTRATION NUMBER: 41,323

TELEPHONE: (617) 498-8284

TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 1461 amino acids

QY 1 PLPPMSGSLKRRPQG 17

|||:|:|:|

Db 420 PLGPDYKGLIRRSPTG 436

RESULT 13

US-08-993-228-10

Sequence 10, Application US/08993228

Patent No. 5976838

GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth

APPLICANT: McCoy, John M.

APPLICANT: LaVallie, Edward R.

APPLICANT: Racie, Lisa A.

APPLICANT: Merberg, David

APPLICANT: Treacy, Maurice

APPLICANT: Spaulding, Vikki

APPLICANT: Agostino, Michael J.

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES

TITLE OF INVENTION: ENCODING THEM

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genetics Institute, Inc.

STREET: 87 CambridgePark Drive

CITY: Cambridge

STATE: MA

COUNTRY: U.S.A.

ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/993,228

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Sprunger, Suzanne A.

REGISTRATION NUMBER: 41,323

TELEPHONE: (617) 498-8284

TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 1461 amino acids

QY 1 PLPPMSGSLKRRPQG 17

|||:|:|:|

Db 420 PLGPDYKGLIRRSPTG 436

Query Match 42.6%; Score 46; DB 4; Length 1124;

Best Local Similarity 47.1%; Pred. No. 75;

Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 PLPPMSGSLKRRPQG 17

|||:|:|:|

Db 420 PLGPDYKGLIRRSPTG 436

US-09-191-786-1

Query Match 42.6%; Score 46; DB 4; Length 1124;

Best Local Similarity 47.1%; Pred. No. 75;

Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 PLPPMSGSLKRRPQG 17

|||:|:~|:|

Db 420 PLGPDYKGLIRRSPTG 436

; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-993-228-10

Query Match 42.6%; Score 46; DB 2; Length 1461;
Best Local Similarity 60.0%; Pred. No. 1e+02;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 LPPEMSGSLKRRPQ 16
:||||| |
Db 272 VPDEMSADLEKRRPE 286

RESULT 14

US-09-006-353A-6
; Sequence 6, Application US/09006353A
; Patent No. 6261801

GENERAL INFORMATION:

; APPLICANT: WEI, YING-FEI
; APPLICANT: YU, GUO-LIANG
; APPLICANT: GENTZ, REINER
; APPLICANT: RUBEN, STEVEN
; TITLE OF INVENTION: TUMOR NECROSIS FACTOR RECEPTOR 5
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: US
; ZIP: 20850

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/006,353A
; FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

; NAME: BROOKES, ANDERS A
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF341
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 415 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-006-353A-6

Query Match 41.7%; Score 45; DB 4; Length 415;
Best Local Similarity 72.7%; Pred. No. 37;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 8 GSKLRRPQGE 18
:||||| |
Db 255 GTLLKRRPGE 265

RESULT 15

US-09-573-986-6

; Sequence 6, Application US/09573986

; Patent No. 6455040

GENERAL INFORMATION:

; APPLICANT: Wei, Ying-Fei

; APPLICANT: Ni, Jian
; APPLICANT: Gentz, Reiner
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: Tumor Necrosis Factor Receptor 5
; FILE REFERENCE: 1488.1280004
; CURRENT APPLICATION NUMBER: US/09/573,986
; CURRENT FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-573-986-6

Query Match 41.7%; Score 45; DB 4; Length 415;
Best Local Similarity 72.7%; Pred. No. 37;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 8 GSKLRRPQGE 18
:||||| |
Db 255 GTLLKRRPGE 265

Search completed: April 8, 2003, 10:53:21
Job time : 3.76611 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2003, 10:51:46 ; Search time 2.14797 Seconds
(without alignments)
569.244 Million cell updates/sec

Title: US-09-917-372-1_COPY_216_235
Perfect score: 108
Sequence: 1 PLPPMSGSLKRRPQGECP 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 248812 segs, 61136040 residues
Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pcp:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pcp:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pcp:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pcp:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pcp:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pcp:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pcp:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pcp:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pcp:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pcp:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pcp:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	108	100.0	399	10	US-09-907-372-1
2	80	74.1	257	10	US-09-948-018-19
3	80	74.1	435	10	US-09-907-372-19
4	80	74.1	435	10	US-09-768-779A-6
5	49	45.4	601	10	US-09-925-301-844
6	48	44.4	930	9	US-10-113-794A-1
7	46	42.6	197	9	US-10-003-211-1
8	46	42.6	348	9	US-09-789-054A-20
9	46	42.6	371	10	US-09-925-300-1681
10	46	42.6	788	9	US-09-978-248-2
11	46	42.6	1124	10	US-09-771-161A-199
12	45	41.7	192	10	US-09-925-302-580
13	45	41.7	264	9	US-10-102-627-67
14	45	41.7	415	10	US-09-826-212-6
15	45	41.7	415	10	US-09-907-372-20
16	45	41.7	415	10	US-09-935-727-8
17	45	41.7	477	9	US-10-047-412A-10
18	45	41.7	543	10	US-09-927-738-2
19	44	40.7	235	9	US-10-050-704-102

20	44	40.7	244	9	US-10-050-704-209
21	44	40.7	477	9	US-10-047-412A-29
22	44	40.7	477	10	US-09-923-556-2
23	44	40.7	477	10	US-09-923-556-6
24	44	40.7	477	10	US-09-987-025-2
25	44	40.7	726	10	US-09-770-689A-4
26	44	40.7	803	10	US-09-770-689A-2
27	44	40.7	881	10	US-09-816-860A-2
28	42	38.9	172	9	US-09-738-626-4747
29	42	38.9	193	10	US-09-813-329-15
30	42	38.9	255	9	US-09-738-626-3575
31	42	38.9	314	9	US-10-097-065-305
32	42	38.9	418	10	US-09-771-161A-218
33	42	38.9	428	9	US-10-125-635A-405
34	42	38.9	570	10	US-09-877-633-13
35	42	38.9	570	10	US-09-877-633-14
36	42	38.9	1265	10	US-09-862-027-80
37	41	38.0	53	10	US-09-864-761-39525
38	41	38.0	201	9	US-10-043-487-323
39	41	38.0	345	9	US-09-738-626-3894
40	41	38.0	487	10	US-09-925-301-910
41	41	38.0	527	9	US-10-022-832-6
42	41	38.0	535	9	US-10-043-487-306
43	41	38.0	793	10	US-09-875-724-4
44	41	38.0	844	10	US-09-875-724-8
45	41	38.0	960	10	US-09-826-508-14

ALIGNMENTS

RESULT 1
US-09-907-372-1
; Sequence 1, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 7497867CDI
US-09-907-372-1

Query Match 100.0%; Score 108; DB 10; Length 399;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 PLPPMSGSLKRRPQGECP 20
DB 216 PLPPMSGSLKRRPQGECP 235
RESULT 2
US-09-948-018-19
; Sequence 19, Application US/09948018
; Patent No. US20020150977A1
; GENERAL INFORMATION:
; APPLICANT: Theill et al
; TITLE OF INVENTION: TNF RECEPTOR-LIKE MOLECULES AND USES THEREOF
; FILE REFERENCE: 01017/37677
; CURRENT APPLICATION NUMBER: US/09/948,018
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/230,191

```
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-948-018-19

Query Match          74.1%; Score 80; DB 10; Length 257;
Best Local Similarity 35.7%; Pred. No. 0.00029;
Matches 20; Conservative 0; Mismatches 0; Indels 36; Gaps 1;

QY 1 PLPPMS-----GSLKKRRPQGECP 20
      |||||
Db 127 PLPPMSGTMLMLAVLLPLAFFLLLATVFCIWKSHPSLCRKLGSLKKRRPQGECP 182

RESULT 3
US-09-907-372-19
; Sequence 19, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 19
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 g339762
US-09-907-372-19

Query Match          74.1%; Score 80; DB 10; Length 435;
Best Local Similarity 35.7%; Pred. No. 0.00052;
Matches 20; Conservative 0; Mismatches 0; Indels 36; Gaps 1;

QY 1 PLPPMS-----GSLKKRRPQGECP 20
      |||||
Db 216 PLPPMSGTMLMLAVLLPLAFFLLLATVFCIWKSHPSLCRKLGSLKKRRPQGECP 271

RESULT 4
US-09-768-779A-6
; Sequence 6, Application US/09768779A
; Patent No. US20020127637A1
; GENERAL INFORMATION:
; APPLICANT: NI, JIAN
; MOORE, PAUL
; TITLE OF INVENTION: HUMAN TUMOR NECROSIS FACTOR
; RECEPTOR-LIKE PROTEIN 8
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/768,779A

; FILING DATE: 25-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/086,582
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: KENLEY K. HOOVER
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PF368PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 435 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-768-779A-6

Query Match          74.1%; Score 80; DB 10; Length 435;
Best Local Similarity 35.7%; Pred. No. 0.00052;
Matches 20; Conservative 0; Mismatches 0; Indels 36; Gaps 1;

QY 1 PLPPMS-----GSLKKRRPQGECP 20
      |||||
Db 216 PLPPMSGTMLMLAVLLPLAFFLLLATVFCIWKSHPSLCRKLGSLKKRRPQGECP 271

RESULT 5
US-09-925-301-844
; Sequence 844, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 844
; LENGTH: 601
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (36)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (64)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (103)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (106)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (152)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (358)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (383)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
```

US-09-925-301-844

Query Match 45.4%; Score 49; DB 10; Length 601;
Best Local Similarity 47.4%; Pred. No. 26;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 LPPEMSGSLKRRPQEGP 20
||| | | :||| |
Db 194 LPPHSSGFLGSKPEGPG 212

RESULT 6

US-10-113-794A-1

; Sequence 1, Application US/10113794A

; Publication No. US20030022202A1

; GENERAL INFORMATION:

; APPLICANT: Flanagan et al.

; TITLE OF INVENTION: B EPHRIN REGULATION OF G-PROTEIN COUPLED

; FILE OF INVENTION: CHEMOATTRACTION

; FILE REFERENCE: 2535/106

; CURRENT APPLICATION NUMBER: US/10/113,794A

; CURRENT FILING DATE: 2002-04-01

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 930

; TYPE: PRT

; ORGANISM: Mus musculus

US-10-113-794A-1

Query Match

44.4%; Score 48; DB 9; Length 930;

Best Local Similarity 50.0%; Pred. No. 59;

Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 PLPPEMSGSLKRRPQEGP 20

| : ||| | : ||| | |

Db 535 VPIPESSASVQKRLPSQEGP 554

RESULT 7

US-10-003-211-1

; Sequence 1, Application US/10003211

; Publication No. US20020197254A1

; GENERAL INFORMATION:

; APPLICANT: Biogen, Inc.

; TITLE OF INVENTION: Soluble Lymphotoxin Beta Receptor and

; TITLE OF INVENTION: Anti-Lymphotoxin Receptor and Ligand Antibodies as

; TITLE OF INVENTION: Therapeutic Agents for the Treatment of Immunological

; FILE REFERENCE: A01305

; CURRENT APPLICATION NUMBER: US/10/003,211

; CURRENT FILING DATE: 2001-10-31

; PRIOR APPLICATION NUMBER: PCT/US97/19436

; PRIOR FILING DATE: 1997-10-24

; PRIOR APPLICATION NUMBER: 60/029,060

; PRIOR FILING DATE: 1996-10-25

; NUMBER OF SEQ ID NOS: 1

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 197

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-003-211-1

Query Match

42.6%; Score 46; DB 9; Length 197;

Best Local Similarity 88.9%; Pred. No. 20;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLPPEMSG 9

|||||

Db 189 PLPPEMSGT 197

RESULT 8

US-09-789-054A-20

; Sequence 20, Application US/09789054A

; Publication No. US20020184659A1

; GENERAL INFORMATION:

; APPLICANT: Allen, Steve

; TITLE OF INVENTION: PLANT GENES ENCODING DRI AND DRAP1, A GLOBAL REPRESSOR COMPLEX

; FILE OF INVENTION: TRANSCRIPTION

; FILE REFERENCE: BB1107 US CIP

; CURRENT APPLICATION NUMBER: US/09/789,054A

; CURRENT FILING DATE: 2001-02-20

; PRIOR APPLICATION NUMBER: 09/485558

; PRIOR FILING DATE: 2000-02-11

; PRIOR APPLICATION NUMBER: PCT/US98/16688

; PRIOR FILING DATE: 1998-08-12

; PRIOR APPLICATION NUMBER: 60/055,865

; PRIOR FILING DATE: 1997-08-15

; NUMBER OF SEQ ID NOS: 69

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 20

; LENGTH: 348

; TYPE: PRT

; ORGANISM: Triticum aestivum

US-09-789-054A-20

Query Match

42.6%; Score 46; DB 9; Length 348;

Best Local Similarity 44.4%; Pred. No. 38;

Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 PPEMSGSLKRRPQEGP 20

| : ||| | : ||| | |

Db 26 PPVPGGTRRRRRPQAQAP 43

RESULT 9

US-09-925-300-1681

; Sequence 1681, Application US/09925300

; Patent No. US20020151681A1

; GENERAL INFORMATION:

; APPLICANT: Craig Rosen,

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

; FILE REFERENCE: PA101

; CURRENT APPLICATION NUMBER: US/09/925,300

; CURRENT FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: PCT/US00/05988

; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 60/124,270

; PRIOR FILING DATE: 1999-03-12

; NUMBER OF SEQ ID NOS: 1890

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1681

; LENGTH: 371

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-925-300-1681

Query Match

42.6%; Score 46; DB 10; Length 371;

Best Local Similarity 52.6%; Pred. No. 41;

Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 LPPEMSGSLKRRPQEGP 20

||| ||| | : | : |

Db 271 LPPTSGSRTKQRLPGQKP 289

RESULT 10

US-09-978-248-2

; Sequence 2, Application US/09978248

; Publication No. US20020197690A1

; GENERAL INFORMATION:

; APPLICANT: Holland, Pamela M
; APPLICANT: Virca, Duke G
; APPLICANT: Bird, Timothy A
; APPLICANT: Garika, Kristen
; TITLE OF INVENTION: GID (GNK INTERACTING DECARBOXYLASE) AND METHODS OF USE
; FILE REFERENCE: 2499-1-001N
; CURRENT APPLICATION NUMBER: US/09/978,248
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/241,324
; PRIOR FILING DATE: 2001-10-18
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 788
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-978-248-2

Query Match 42.6%; Score 46; DB 9; Length 788;
Best Local Similarity 52.6%; Pred. No. 96;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 LPPEMSGSLKRRPQGCP 20
Db 688 LPPTSGSRKORLPQGRP 706

RESULT 11

US-09-771-161A-199
; Sequence 199, Application US/09771161A
; Patent No. US2002011081A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 199
; LENGTH: 1124
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-771-161A-199

Query Match 42.6%; Score 46; DB 10; Length 1124;
Best Local Similarity 47.1%; Pred. No. 1.4e+02;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 PLPPEMSGSLKRRPQG 17
Db 420 PLGPDYRGCLIRSPGT 436

RESULT 12

US-09-925-302-580
; Sequence 580, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270

; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 580
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-302-580

Query Match 41.7%; Score 45; DB 10; Length 192;
Best Local Similarity 50.0%; Pred. No. 27;
Matches 11; Conservative 4; Mismatches 3; Indels 4; Gaps 1;

QY 2 LPPEMSGSLKRR----PQEGG 19
Db 162 LCPQVSGSIHKRKIHFFPQGWG 183

RESULT 13

US-10-102-627-67
; Sequence 67, Application US/10102627
; Publication No. US20030054377A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ24C1
; CURRENT APPLICATION NUMBER: US/10/102,627
; CURRENT FILING DATE: 2002-03-22
; NUMBER OF SEQ ID NOS: 110
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 67
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (111)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-102-627-67

Query Match 41.7%; Score 45; DB 9; Length 264;
Best Local Similarity 53.3%; Pred. No. 39;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 PLPPEMSGSLKRRP 15
Db 223 PLPQDQTESLLRRQP 237

RESULT 14

US-09-826-212-6
; Sequence 6, Application US/09826212
; Patent No. US20010021516A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Ying-Fei
; APPLICANT: Gentz, Reiner
; APPLICANT: Ruben, Steven
; APPLICANT: Ni, Jian
; TITLE OF INVENTION: Tumor Necrosis Factor Receptor 5
; FILE REFERENCE: 1488.1280006
; CURRENT APPLICATION NUMBER: US/09/826,212
; CURRENT FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-826-212-6

Query Match 41.7%; Score 45; DB 10; Length 415;
Best Local Similarity 72.7%; Pred. No. 65;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 8 GSKLKRPEGE 18
|:||||| |:|
Db 255 GTLLKRPEGE 265

RESULT 15
US-09-907-372-20
; Sequence 20, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Warren, Bridget A.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 20
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 g600223
US-09-907-372-20

Query Match 41.7%; Score 45; DB 10; Length 415;
Best Local Similarity 72.7%; Pred. No. 65;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 8 GSKLKRPEGE 18
|:||||| |:|
Db 255 GTLLKRPEGE 265

Search completed: April 8, 2003, 11:03:51
Job time : 3.14797 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:05:15 ; Search time 344.98 seconds
(without alignments)
12938.300 Million cell updates/sec

Title: US-09-917-372-2
Perfect score: 1982
Sequence: 1 gccccgcgcagctcgtc.....ggtatggggagaggtttgg 1982

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*

2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*

3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*

4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*

5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*

6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*

7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*

8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*

9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*

10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*

11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*

12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*

13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*

14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*

15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*

16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*

17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*

18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*

19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*

20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*

21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*

22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*

23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*

24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1809.2	91.3	2136	24	Gene #2135 used to
2	1809.2	91.3	2136	24	Human benign prost
3	1809.2	91.3	2136	24	Breast cancer rela
4	1809.2	91.3	2136	24	Stomach cancer rel
5	1173.4	59.2	1594	24	Human ovarian anti
6	444.6	22.4	527	24	Human colon cancer
7	433.6	21.9	481	24	Human ovarian can
8	386.4	19.5	529	24	Human ovarian can
9	382	19.3	516	24	Human ovarian can

c	10	253.8	12.8	281	24	ABL82623	Human ovarian can
	11	251.2	12.7	289	24	ABL82688	Human ovarian can
	12	72	3.6	1557	15	AAQ45225	Sequence encoding
	13	72	3.6	1557	22	AAQ90446	TNFR:Fc fusion pro
	14	72	3.6	1840	12	AAQ10990	Partial sequence o
	15	72	3.6	1641	14	AAQ49931	TNF-R CDNA. Homo
	16	72	3.6	2393	12	AAQ10907	40kD TNF inhibitor
	17	72	3.6	2394	22	AAQ83951	Human 40 kDa TNF i
	18	72	3.6	3683	24	ABK83977	Human CDNA differe
	19	72	3.6	3683	24	ABK65877	Lung cancer relate
	20	72	3.6	3683	24	ABK33465	Human TNF receptor
	21	72	3.6	3683	24	ABK33466	Human TNF receptor
	22	72	3.6	3683	24	ABK33467	Human TNF receptor
	23	70.4	3.6	2224	16	AAQ89544	p75 Tumour Necrosi
	24	69.2	3.5	1641	15	AAQ45224	Sequence encoding
	25	62.6	3.2	691	16	AAQ05443	BamTP delta53 nerv
	26	62.6	3.2	705	19	AAV41549	Human soluble tumo
	27	62.6	3.2	705	19	AAV19802	Tumour necrosis fa
	28	62.6	3.2	705	20	AAV81733	Tumour necrosis in
	29	62.6	3.2	705	21	AAZ45759	DNA encoding a K10
	30	62.6	3.2	705	21	AAZ45760	DNA encoding a K12
	31	62.6	3.2	705	21	AAZ45761	DNA encoding a K10
	32	62.6	3.2	705	21	AAZ45762	Wild type N-termina
	33	62.4	3.1	927	19	AAV44852	Herpesvirus entry
	34	62.4	3.1	927	20	AAZ87265	CDNA clone encodin
	35	62.4	3.1	1049	21	AAA46931	CDNA encoding nove
	36	62.4	3.1	1049	21	AAZ49727	Human PRO509 CDNA
	37	62.4	3.1	1596	21	AAZ94196	Soluble herpesviri
	38	62.4	3.1	1704	21	AAA28135	Human TR2 receptor
	39	62.4	3.1	1724	18	AAZ51737	Human herpes simpl
	40	62.4	3.1	1724	20	AAZ25512	Herpes virus entry
	41	62.4	3.1	1724	24	ABK84410	Human CDNA differe
	42	62.4	3.1	1815	22	AAZ03021	Human diagnostic a
	43	62.4	3.1	1834	21	AAZ94198	Membrane-bound her
	44	62.4	3.1	1929	21	AAZ94195	Soluble herpesviri
	45	62.4	3.1	2313	21	AAZ94197	Soluble herpesviri

ALIGNMENTS

RESULT 1	ABN95637	
ID	ABN95637 standard; DNA; 2136 BP.	
XX	ABN95637;	
AC	ABN95637;	
DT	13-AUG-2002 (first entry)	
XX		
DE	Gene #2135 used to diagnose liver cancer.	
XX		
KW	Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;	
KW	metastatic liver tumour; cytostatic; expression profile; disease state;	
KW	disease progression; drug toxicity; drug efficacy; drug metabolism.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200229103-A2.	
XX		
PD	11-APR-2002.	
XX		
PF	02-OCT-2001; 2001WO-US30589.	
XX		
PR	02-OCT-2000; 2000US-237054P.	
XX	(GENE-) GENE LOGIC INC.	
PA		
PI	Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;	
XX		
DR	WPI; 2002-426119/45.	
XX		
PT	Diagnosing and detecting the progression of liver cancer,	
PT	hepatocellular carcinoma or metastatic liver tumor in a patient,	

PT involves detecting the level of expression of two or more genes in a
 PS liver tissue sample
 XX
 PS Claim 1; SEQ ID NO 2135; 298pp; English.
 XX
 CC The invention relates to a novel method for diagnosing and detecting the
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 CC tumour in a patient, and differentiating metastatic liver cancer from
 CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN93503-ABN97455 in a
 CC tissue sample. The method of the invention has hepatotropic, and
 CC cytostatic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as
 CC markers that can be used to monitor disease states, disease progression,
 CC drug toxicity, drug efficacy and drug metabolism.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 2136 BP; 446 A; 706 C; 608 G; 376 T; 0 other;

Query Match 91.3%; Score 1809.2; DB 24; Length 2136;
 Best Local Similarity 94.6%; Pred. No. 0;
 Matches 1929; Conservative 0; Mismatches 3; Indels 108; Gaps 1;

QY 51 GCCCTGAGGCGCGCTGCGCGCTCCCGCCCTGGGTGCACATCGCCCTGAGTCCCG 110
 DB 1 GCCCTGAGGCGCGCTGCGCGCTCCCGCCCTGGGTGCACATCGCCCTGAGTCCCG 60
 QY 111 TCCAGGCTCTGGGCTGGGAGCGCGCGCCGCTGCCAGAGCTGGGCGCTCTGCG 170
 DB 61 TCCAGGCTCTGGGCTGGGAGCGCGCGCCGCTGCCAGAGCTGGGCGCTCTGCG 120
 QY 171 CTTCTCCAGGCGCGCGCTGCTGCGCGCTGCGCGCTGCGCGCTGCTGCGCT 230
 DB 121 CTTCTCCAGGCGCGCGCTGCTGCGCGCTGCGCGCTGCGCGCTGCTGCGCT 180
 QY 231 TGGGCGACCTCTGCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCTGCGCGCT 290
 DB 181 TGGGCGACCTCTGCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCTGCGCGCT 240
 QY 291 CTGGCAGCATCGAGCGCGCGCTGCTGCGCGCTGCGCGCTGCGCGCTGCTGCGCG 350
 DB 241 CTGGCAGCATCGAGCGCGCGCTGCTGCGCGCTGCGCGCTGCGCGCTGCTGCGCG 300
 QY 351 GACCAGAAAGGAATACTATGAGCGCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCGCG 410
 DB 301 GACCAGAAAGGAATACTATGAGCGCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCGCG 360
 QY 411 GGCACCTATGCTCAGCTAATGAGCGCGCGCTGCGCGCGCTGCTGCGCGCTGCTGCGCG 470
 DB 361 GGCACCTATGCTCAGCTAATGAGCGCGCGCTGCGCGCGCTGCTGCGCGCTGCTGCGCG 420
 QY 471 GAGAAATCTTACACAGGAGCTGGAATCTGACCATCTGCGAGCTGTCGCGCGCTGCT 530
 DB 421 GAGAAATCTTACACAGGAGCTGGAATCTGACCATCTGCGAGCTGTCGCGCGCTGCT 480
 QY 531 GACCCAGTGTGGGCTGAGGAGATGCGCGCTGCGAGCAAGCAAGCAAGCAAGCAAGCAAG 590
 DB 481 GACCCAGTGTGGGCTGAGGAGATGCGCGCTGCGAGCAAGCAAGCAAGCAAGCAAGCAAG 540
 QY 591 CGTGTCCAGCGGGAATGTTCTGTGCTGCTGCGCGCGCTGCGAGTGTACACACTGCGAGCTA 650
 DB 541 CGTGTCCAGCGGGAATGTTCTGTGCTGCTGCGCGCGCTGCGAGTGTACACACTGCGAGCTA 600
 QY 651 CTTTCTGACTGCGCGCTGCGAGCTGAGCGAGCTCAAGATGAAGTTGGAGAGGTTAAC 710
 DB 601 CTTTCTGACTGCGCGCTGCGAGCTGAGCGAGCTCAAGATGAAGTTGGAGAGGTTAAC 660
 QY 711 AACCACTGCGCGCTGCGAGGAGGAGCTTCCAGAAATACCTCTCCCGCAGCGCGCG 770
 DB 771 AACCACTGCGCGCTGCGAGGAGGAGCTTCCAGAAATACCTCTCCCGCAGCGCGCG 720

DB 661 AACCACTGCGCGCTGCGAGGAGGAGCTTCCAGAAATACCTCTCCCGCAGCGCGCG 720
 QY 771 TGGCAGCGCGCGAGCTGTGAGCAACCAAGGTCTGTGGAGGAGCTTCCAGGCACTGCC 830
 DB 721 TGGCAGCGCGCGAGCTGTGAGCAACCAAGGTCTGTGGAGGAGCTTCCAGGCACTGCC 780
 QY 831 CAGTCCGACACACCTGCAAAATATCCATTAGAGCCACTGCCCGCAGAGATGTCA 884
 DB 781 CAGTCCGACACACCTGCAAAATATCCATTAGAGCCACTGCCCGCAGAGATGTCA 840
 QY 885 ----- 884
 DB 841 ATGCTGATGTGGCGTCTGCTGCCACTGGCGCTTCTTCTGCTCTGCCACCGTCTTC 900
 QY 885 -----GGATCGCTGCTCAAGAGG 902
 DB 901 TCCTGCTATCTGGAAGAGCCACCTTCTCTGCGAGAAACTGGGATCGCTGCTCAAGAGG 960
 QY 903 CGTCCGAGGAGGAGGAGCCCAATCTCTGAGCTGGAGCTGGGAGCTTCCGAGGCCAT 962
 DB 961 CGTCCGAGGAGGAGGAGCCCAATCTCTGAGCTGGAGCTGGGAGCTTCCGAGGCCAT 1020
 QY 963 CCATACCTTCTGACTTGTGTACAGCCACTGCTACCCATTTCTGGAGATGTTTCCCGCAGTA 1022
 DB 1021 CCATACCTTCTGACTTGTGTACAGCCACTGCTACCCATTTCTGGAGATGTTTCCCGCAGTA 1080
 QY 1023 TCCACTGGGCTCCCGCAGCGCCAGTTTGGAGCGAGGGTGGCGCAACAGCAGAGTCT 1082
 DB 1081 TCCACTGGGCTCCCGCAGCGCCAGTTTGGAGCGAGGGTGGCGCAACAGCAGAGTCT 1140
 QY 1083 CTGGACCTGACAGGAGCGCGAGTTGGAACCCGGGAGCAGAGCCAGGTGGGCCACGGT 1142
 DB 1141 CTGGACCTGACAGGAGCGCGAGTTGGAACCCGGGAGCAGAGCCAGGTGGGCCACGGT 1200
 QY 1143 ACCAATGGCATTCATGTCACCGCGGTCTATGACTATCTAGTGCACACATCTACATCTAC 1202
 DB 1201 ACCAATGGCATTCATGTCACCGCGGTCTATGACTATCTAGTGCACACATCTACATCTAC 1260
 QY 1203 AATGGACCATCTAGTGGGGGAGCACCAGGCTCTGGAGACCTCCAGTACCCCGAGACCT 1262
 DB 1261 AATGGACCATCTAGTGGGGGAGCACCAGGCTCTGGAGACCTCCAGTACCCCGAGACCT 1320
 QY 1263 CCATACCCCATTCGGAAGAGGGGAGCCCTGGCGCTCCCGGCTCTACACCCACAG 1322
 DB 1321 CCATACCCCATTCGGAAGAGGGGAGCCCTGGCGCTCCCGGCTCTACACCCACAG 1380
 QY 1323 GAAGATGGCAAGGCTTGGCACCTAGCGAGACAGACACTGTGTGCGCACACCCCTCTAAC 1382
 DB 1381 GAAGATGGCAAGGCTTGGCACCTAGCGAGACAGACACTGTGTGCGCACACCCCTCTAAC 1440
 QY 1383 AGGGGCCAAGGAACCAATTTATCACCATTGACTGACTGAGTGTGAGAAAGGAGCAAGA 1442
 DB 1441 AGGGGCCAAGGAACCAATTTATCACCATTGACTGAGGAGTGTGAGAAAGGAGCAAGA 1500
 QY 1443 AGGGGGCACAAGGACCTTCTCCCTTGGCTGGCTGCCCTGCCACGTGGGATTCACAGGG 1502
 DB 1501 AGGGGGCACAAGGACCTTCTCCCTTGGCTGGCTGCCCTGCCACGTGGGATTCACAGGG 1560
 QY 1503 GCCTGAGTAGGGCCCGGGGAAGCAGAGCCCTTAAGGGATTAAAGCTCAGACACCTCTCAGA 1562
 DB 1561 GCCTGAGTAGGGCCCGGGGAAGCAGAGCCCTTAAGGGATTAAAGCTCAGACACCTCTCAGA 1620
 QY 1563 GCAGGTGGGCACTGGCTGGGTACGGTCCCTCCACAGGACTCTCCCTACTGCTGAGCAA 1622
 DB 1621 GCAGGTGGGCACTGGCTGGGTACGGTCCCTCCACAGGACTCTCCCTACTGCTGAGCAA 1680
 QY 1623 ACCTGAGGCGCTCCCGGAGAGCCACCCCTGGGGCTGCTCAGGCTCAGGACCGGAC 1682
 DB 1681 ACCTGAGGCGCTCCCGGAGAGCCACCCCTGGGGCTGCTCAGGCTCAGGACCGGAC 1740
 QY 1683 AGGCACTCATACCAACTGCTGCCACTACAGCACCCGACCGGAGCAGGACCGGAG 1742
 DB 1741 AGGCACTCATACCAACTGCTGCCACTACAGCACCCGACCGGAGCAGGACCGGAG 1800

Db 901 TCCTGCATCTGGAAGAGCCACCCCTTCTCTCCAGAAACTGGGATCGCTGCTCAAGAGG 960
QY 903 CGTCCGAGGAGAGGGACCCCAATCCTGTAGCTGGAAGCTGGAGCCCTCCGGAAGGCCCAT 962
Db 961 CGTCCGAGGAGAGGGACCCCAATCCTGTAGCTGGAAGCTGGAGCCCTCCGGAAGGCCCAT 1020
QY 963 CCATATCTCCCTGACTGTGTACAGCACTGCTTACCACTTCTGGAGATGTTTCCCCAGTA 1022
Db 1021 CCATATCTCCCTGACTGTGTACAGCACTGCTTACCACTTCTGGAGATGTTTCCCCAGTA 1080
QY 1023 TCCACTGGGCTCCCGCAGCCCCAGTTTTTGGAGGAGGGGTCCGCAACAGCAGAGTCTCT 1082
Db 1081 TCCACTGGGCTCCCGCAGCCCCAGTTTTTGGAGGAGGGGTCCGCAACAGCAGAGTCTCT 1140
QY 1083 CTGGACCTACACAGGAGCGCAGTTGGAAACCCGGGAGCAGAGCAGGTGGCCACCGGT 1142
Db 1141 CTGGACCTACACAGGAGCGCAGTTGGAAACCCGGGAGCAGAGCAGGTGGCCACCGGT 1200
QY 1143 ACCAATGGCATTCATGTACCGGGCGGGTCTATGACTATCATCTGGCAACATCTACATCTAC 1202
Db 1201 ACCAATGGCATTCATGTACCGGGCGGGTCTATGACTATCATCTGGCAACATCTACATCTAC 1260
QY 1203 AATGACCACTACTGGGGGACCCACCGGTCTCTGGAGACCTCCAGCTACCCCGAACCT 1262
Db 1261 AATGACCACTACTGGGGGACCCACCGGTCTCTGGAGACCTCCAGCTACCCCGAACCT 1320
QY 1263 CCATACCCCATTCCTCGAAGAGGGGACCTGTGCGCTCCCGGCTCTCTACACCCACCCAG 1322
Db 1321 CCATACCCCATTCCTCGAAGAGGGGACCTGTGCGCTCCCGGCTCTCTACACCCACCCAG 1380
QY 1323 GAAGATGCAAGGCTTGGCACTACGGAGACAGAGCACTGTGGTGCACACCCCTCTAAC 1382
Db 1381 GAAGATGCAAGGCTTGGCACTACGGAGACAGAGCACTGTGGTGCACACCCCTCTAAC 1440
QY 1383 AGGGGCCCCAAGGAACCAATTTATACCCATGACTGACTGAGTCTCAGAAAGGCGAGAAGA 1442
Db 1441 AGGGGCCCCAAGGAACCAATTTATACCCATGACTGACTGAGTCTCAGAAAGGCGAGAAGA 1500
QY 1443 AGGGGGGACAAAGGCACCTTCTCCCTTTAGGCTGCCCTGCCACGTGGGATTCACAGGG 1502
Db 1501 AGGGGGGACAAAGGCACCTTCTCCCTTTAGGCTGCCCTGCCACGTGGGATTCACAGGG 1560
QY 1503 GCCTGAGTAGGGCCCGGGAGCAGACAGCCCTAAGGGATTAGGCTCAGACACCTCTGAGA 1562
Db 1561 GCCTGAGTAGGGCCCGGGAGCAGACAGCCCTAAGGGATTAGGCTCAGACACCTCTGAGA 1620
QY 1563 GCAGTGGGCACTGGCTGGGTACGGTGCCTCCACAGGACTCTCCCTACTGCTGCTGAGCAA 1622
Db 1621 GCAGTGGGCACTGGCTGGGTACGGTGCCTCCACAGGACTCTCCCTACTGCTGCTGAGCAA 1680
QY 1623 ACCTGAGGCTCCCGGACAGACCCACCCACCCCTGGGCTGTGCTCAGGCTCAGGCACGGAC 1682
Db 1681 ACCTGAGGCTCCCGGACAGACCCACCCACCCCTGGGCTGTGCTCAGGCTCAGGCACGGAC 1740
QY 1683 AGGGACATGATACCACTGTGCCCACTACAGACGCGGACCGGAGCAGCGCACCGAG 1742
Db 1741 AGGGACATGATACCACTGTGCCCACTACAGACGCGGACCGGAGCAGCGCACCGAG 1800
QY 1743 GGAGCGCCACACGGTCACTGCAAGGACGTCACGGGCCCTCTAAAGGATTCGTTGTC 1802
Db 1801 GGAGCGCCACACGGTCACTGCAAGGACGTCACGGGCCCTCTAAAGGATTCGTTGTC 1860
QY 1803 TCATCCCCAAGCTTCAGAGACCCCTTTGGGGTTCCACACTTACGTTGGACTGAGTAGACC 1862
Db 1861 TCATCCCCAAGCTTCAGAGACCCCTTTGGGGTTCCACACTTACGTTGGACTGAGTAGACC 1920
QY 1863 CTGCATGAAGTGAATTTATAGGAGGAGCGTCTCTCCCTCCCTCTCTAGAGGAGAGAA 1922
Db 1921 CTGCATGAAGTGAATTTATAGGAGGAGCGTCTCTCCCTCCCTCTCTAGAGGAGAGAA 1980
QY 1923 AGGGAGTCAATTAACAACCTAGGGGTTGGGTAGGATTCCTAGGTATGGGAAGAGTTTGG 1982
Db 1981 AGGGAGTCAATTAACAACCTAGGGGTTGGGTAGGATTCCTAGGTATGGGAAGAGTTTGG 2040

RESULT 3

ABL64078

ID ABL64078 standard; DNA; 2136 Bp.

XX ABL64078;

XX ABL64078;

XX 15-MAY-2002 (first entry)

XX Breast cancer related gene sequence SEQ ID NO:2415.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;

XX stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;

XX cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;

XX gene; ds.

XX Homo sapiens.

XX WO200194629-A2.

XX 13-DEC-2001.

XX 30-MAY-2001; 2001WO-US10838.

XX 05-JUN-2000; 2000US-209473P.

XX 05-JUN-2000; 2000US-209531P.

XX 18-SEP-2000; 2000US-233133P.

XX 18-SEP-2000; 2000US-233617P.

XX 20-SEP-2000; 2000US-234009P.

XX 20-SEP-2000; 2000US-234034P.

XX 20-SEP-2000; 2000US-234052P.

XX 22-SEP-2000; 2000US-234509P.

XX 22-SEP-2000; 2000US-234567P.

XX 25-SEP-2000; 2000US-234923P.

XX 25-SEP-2000; 2000US-234924P.

XX 25-SEP-2000; 2000US-235077P.

XX 25-SEP-2000; 2000US-235082P.

XX 25-SEP-2000; 2000US-235134P.

XX 25-SEP-2000; 2000US-235280P.

XX 26-SEP-2000; 2000US-235637P.

XX 26-SEP-2000; 2000US-235638P.

XX 27-SEP-2000; 2000US-235711P.

XX 27-SEP-2000; 2000US-235720P.

XX 27-SEP-2000; 2000US-235840P.

XX 27-SEP-2000; 2000US-235863P.

XX 28-SEP-2000; 2000US-236028P.

XX 28-SEP-2000; 2000US-236032P.

XX 28-SEP-2000; 2000US-236033P.

XX 28-SEP-2000; 2000US-236034P.

XX 28-SEP-2000; 2000US-236109P.

XX 28-SEP-2000; 2000US-236111P.

XX 29-SEP-2000; 2000US-236842P.

XX 29-SEP-2000; 2000US-236891P.

XX 02-OCT-2000; 2000US-237172P.

XX 02-OCT-2000; 2000US-237173P.

XX 02-OCT-2000; 2000US-237278P.

XX 02-OCT-2000; 2000US-237294P.

XX 02-OCT-2000; 2000US-237295P.

XX 02-OCT-2000; 2000US-237316P.

XX 03-OCT-2000; 2000US-237425P.

XX 03-OCT-2000; 2000US-237598P.

XX 03-OCT-2000; 2000US-237604P.

XX 03-OCT-2000; 2000US-237606P.

XX 03-OCT-2000; 2000US-237608P.

XX 01-NOV-2000; 2000US-244867P.

XX 01-NOV-2000; 2000US-245084P.

XX (AVAL-) AVALON PHARM.

XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;

XX Soppet DR, Weaver Z;

XX

DR WPI: 2002-188264/24.
XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
XX Claim 1; SEQ ID 2415; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilms' tumour.
XX
SQ Sequence 2136 BP; 446 A; 706 C; 608 G; 376 T; 0 other;

Query Watch 91.3%; Score 1809.2; DB 24; Length 2136;
Best Local Similarity 94.6%; Pred. No. 0;
Matches 1929; Conservative 0; Mismatches 3; Indels 108; Gaps 1;

QY 51 GCCTCGAGCCCGCTGCGCTCCCGCCCTGGGTGCACATCGGCCCTGAGTCCCG 110
DB 1 GCCTCGAGCCCGCTGCGCTCCCGCCCTGGGTGCACATCGGCCCTGAGTCCCG 60
QY 111 TCCAGGCTCTGGGCTCGGCGACCGCCGACCGCTGCCAGAGCTCGGGCTCCTGC 170
DB 61 TCCAGGCTCTGGGCTCGGCGACCGCCGACCGCTGCCAGAGCTCGGGCTCCTGC 120
QY 171 CTTCTCCAGGCCCGCCAGCTTCTGGCGCCCTGGCGAGTGGCGCGCATGCTCGCT 230
DB 121 CTTCTCCAGGCCCGCCAGCTTCTGGCGCCCTGGCGAGTGGCGCGCATGCTCGCT 180
QY 231 TGGCCACCTCTGCCCCCGGCTTGGCTGGGGCCCTTGTGCTGGCCCTTTCGGGCTC 290
DB 181 TGGCCACCTCTGCCCCCGGCTTGGCTGGGGCCCTTGTGCTGGCGCTTTCGGGCTC 240
QY 291 CTGGCAGATCGAGCCCGCCAGCGGTGCTTCATATGCTGCGGAGAACAGACTCGAG 350
DB 241 CTGGCAGATCGAGCCCGCCAGCGGTGCTTCATATGCTGCGGAGAACAGACTCGAG 300
QY 351 GACCCAGAAAGAAATACTATGAGCCCGCCAGCCGCTATGCTGCTCCGCTCCCGCCA 410
DB 301 GACCCAGAAAGAAATACTATGAGCCCGCCAGCCGCTATGCTGCTCCGCTCCCGCCA 360
QY 411 GGCACCTATGCTCAGCTAAATAGCCGATCCGGGACACAGTTGTGCGCACATGTGCC 470
DB 361 GGCACCTATGCTCAGCTAAATAGCCGATCCGGGACACAGTTGTGCGCACATGTGCC 420
QY 471 GAGAAATCTACAAAGAGACTGGAATACCTGACCATCTGCCAGTGTGCCGCCCTGT 530
DB 421 GAGAAATCTACAAAGAGACTGGAATACCTGACCATCTGCCAGTGTGCCGCCCTGT 480
QY 531 GACCCAGTATGGCCCTCGAGGAGATTGCCCTTGACACAGCAACAGGAGCCAGTGC 590
DB 481 GACCCAGTATGGCCCTCGAGGAGATTGCCCTTGACACAGCAACAGGAGCCAGTGC 540
QY 591 CGCTGCAGCCGGGAATGTTCTGCTGCTGGGCCCTCGAGTGTACACACTCGGAGCTA 650
DB 541 CGCTGCAGCCGGGAATGTTCTGCTGCTGGGCCCTCGAGTGTACACACTCGGAGCTA 600
QY 651 CTTTCTGACTGCGCCCTGGCACTGAAGCCGAGCTCAAGATGAAGTTGGGAAGGTTAAC 710

DB 601 CTTTCTGACTGCGCGCTGCACTGAAGCGAGCTCAAGATGAAGTTGGGAAGGTTAAC 660
QY 711 AACCACTGGTCCCTTCCGAAAGGAGGCACTTCCAGAAATACCTTCCCGCCAGCGCCGC 770
DB 661 AACCACTGGTCCCTTCCGAAAGGAGGCACTTCCAGAAATACCTTCCCGCCAGCGCCGC 720
QY 771 TGGCAGCCCGCACACAGGCTGTGAGAACCAAGGCTTGGTGGAGGCACTCCAGGCACTGCC 830
DB 721 TGGCAGCCCGCACACAGGCTGTGAGAACCAAGGCTTGGTGGAGGCACTCCAGGCACTGCC 780
QY 831 CAGTCCGACACAACTTGCAGAAATCCATTAGAGCCACTGCCCGCCAGAGATGTCA 884
DB 781 CAGTCCGACACAACTTGCAGAAATCCATTAGAGCCACTGCCCGCCAGAGATGTCA 840
QY 885 ----- 884
DB 841 ATGCTGATGCTGGCCGTTCTGCTGCACTGGCCCTTCTTCTGCTCCTTGGCACCGTCTTC 900
QY 885 -----GGATCGCTGCTCAAGAGG 902
DB 901 TCTGCTGATGGAAGAGCCACCTTCTCTGTCAGGAAACTGGGATCGCTTCAAGAGG 960
QY 903 CGTCCCGAGGAGAGGAGCCCAATCCTGTAGCTGGAAGCTGGGAGCTCCGAAAGGCCAT 962
DB 961 CGTCCCGAGGAGAGGAGCCCAATCCTGTAGCTGGAAGCTGGGAGCTCCGAAAGGCCAT 1020
QY 963 CCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022
DB 1021 CCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 1023 TCCACATGGGCTCCCGCAGCCCGCTTGGAGGAGGCTGCGCAACAGAGAGTCT 1082
DB 1081 TCCACATGGGCTCCCGCAGCCCGCTTGGAGGAGGCTGCGCAACAGAGAGTCT 1140
QY 1083 CTGGACCTGACACAGGAGCGCGAGTTGGAAACCCCGGGAGCAGAGGCTGGCCCGCAGGT 1142
DB 1141 CTGGACCTGACACAGGAGCGCGAGTTGGAAACCCCGGGAGCAGAGGCTGGCCCGCAGGT 1200
QY 1143 ACCAATGGCATTCATGCTACCCGGGGCTTATGACTATCTACTGGCAACATCTACATCTAC 1202
DB 1201 ACCAATGGCATTCATGCTACCCGGGGCTTATGACTATCTACTGGCAACATCTACATCTAC 1260
QY 1203 AATGGACCACTACTGGGGGACCCAGCGGTCTCTGGAGACCTCCAGCTACCCCGCAACCT 1262
DB 1261 AATGGACCACTACTGGGGGACCCAGCGGTCTCTGGAGACCTCCAGCTACCCCGCAACCT 1320
QY 1263 CCATACCCATTTCCGAAAGAGGCGGCTTGGCCCTTCCCGGGCTCTCTACACCCCGCAG 1322
DB 1321 CCATACCCATTTCCGAAAGAGGCGGCTTGGCCCTTCCCGGGCTCTCTACACCCCGCAG 1380
QY 1323 GAAGATGGCAAGGCTTGGCACCTAGCGGAGACAGACCTGCTGGTCCACACCTCTAAC 1382
DB 1381 GAAGATGGCAAGGCTTGGCACCTAGCGGAGACAGACCTGCTGGTCCACACCTCTAAC 1440
QY 1383 AGGGGCCCAAGAACCAATTTATCACCATGACTGACTGAGTCTGAGAAAAGGCAAGA 1442
DB 1441 AGGGGCCCAAGAACCAATTTATCACCATGACTGAGGAGTCTGAGAAAAGGCAAGA 1500
QY 1443 AGGGGGCACAAAGGCACTTCTCCCTTGGAGTGGCTGCCCTGCCACAGTGGGATTCACAGG 1502
DB 1501 AGGGGGCACAAAGGCACTTCTCCCTTGGAGTGGCTGCCCTGCCACAGTGGGATTCACAGG 1560
QY 1503 GCCTGAGTAGGGCCCGGGGAGCAGAGCCCTTAAGGATTAAGGCTCAGACACCTCTGAGA 1562
DB 1561 GCCTGAGTAGGGCCCGGGGAGCAGAGCCCTTAAGGATTAAGGCTCAGACACCTCTGAGA 1620
QY 1563 GCAGGTGGGCACCTGGGTGAGTGGCTTCCCTCCACAGGACTCTCCCTACTGCTGAGCAA 1622
DB 1621 GCAGGTGGGCACCTGGGTGAGTGGCTTCCCTCCACAGGACTCTCCCTACTGCTGAGCAA 1680
QY 1623 ACCTGAGGCTCCTCCGCGAGACCCACCCCTGGGGCTGCTCAGCTCAGGCAAGGAC 1682

QY	351	GACCAGAAAAGAACTACTATGAGCCCCAGCACGCACTCTGCTGCTCCCGCTGCCCGGCCA	410
Db	301		
		GACCAGAAAAGAACTACTATGAGCCCCAGCACGCACTCTGCTGCTCCCGCTGCCCGGCCA	360
QY	411	GGCACCTATGTCCTCAGCTAAATGTAGCCGATCCGGGACACAGTTTGTGCCACATGTGCC	470
Db	361		
		GGCACCTATGTCCTCAGCTAAATGTAGCCGATCCGGGACACAGTTTGTGCCACATGTGCC	420
QY	471	GAGAATTCTCTACAACGAGCACTTGAACCTACCTTGACCATCTGCCAGCTGTGCCGCCCTGT	530
Db	421		
		GAGAATTCTCTACAACGAGCACTTGAACCTACCTTGACCATCTGCCAGCTGTGCCGCCCTGT	480
QY	531	GACCCAGTATGGGCGCTCGAGGAGATTGGCCCGCTGCACAACAACGAAGACCCAGTGC	590
Db	481		
		GACCCAGTATGGGCGCTCGAGGAGATTGGCCCGCTGCACAACAACGAAGACCCAGTGC	540
QY	591	CGCTGCCAGCGGGGAATGTTCTGCTGCTGCTGGGCCCTCGAGTGTACACACTGGGAGCTA	650
Db	541		
		CGCTGCCAGCGGGGAATGTTCTGCTGCTGCTGGGCCCTCGAGTGTACACACTGGGAGCTA	600
QY	651	CTTTCTGACTGCCCGCTGGCACTGAAGCCGAGCTCAAGATGAAGTTGGGAAGGGTAAC	710
Db	601		
		CTTTCTGACTGCCCGCTGGCACTGAAGCCGAGCTCAAGATGAAGTTGGGAAGGGTAAC	660
QY	711	AACCACTGGTCCCTCGAAGCGAGGGCACTTCCAGATACTCTCTCCCGCAGCGCCCGC	770
Db	661		
		AACCACTGGTCCCTCGAAGCGAGGGCACTTCCAGATACTCTCTCCCGCAGCGCCCGC	720
QY	771	TGCCAGCCCCACACAGAGTGTACAGAACCAAGTCTGTGTGAGGAGCACTCCAGGCACTGCC	830
Db	721		
		TGCCAGCCCCACACAGAGTGTACAGAACCAAGTCTGTGTGAGGAGCACTCCAGGCACTGCC	780
QY	831	CAGTCCGACACAACCTGCAAAAATCATTAGAGCCACTGCCCGCCAGAGATGTCA-	884
Db	781		
		CAGTCCGACACAACCTGCAAAAATCCATTAGAGCCACTGCCCGCCAGAGATGTCAAGAAC	840
QY	885	-----	884
Db	841	ATGCTGATGCTGGCGGTCTGCTGCCACTGGCGCTTCTTTCTGCTCCTTGCCACCGTCTTC	900
QY	885	-----	902
Db	901		
		TCCTGCATCTGGAGAGCCACCCCTTCTCTCTGCAGGAACCTGGGATCGCTCTCAAGAGG	960
QY	903	CGTCCGAGGAGAGGAGGCCCAATCTGTAGCTGGAAGCTGGGAGCCTCCGAAGGCCCAT	962
Db	961		
		CGTCCGAGGAGAGGAGGCCCAATCTGTAGCTGGAAGCTGGGAGCCTCCGAAGGCCCAT	1020
QY	963	CCATACTTCCCTGACTTGTGTACAGCCACTGTACCCATTTCTGGAGATGTTTCCCCAGTA	1022
Db	1021		
		CCATACTTCCCTGACTTGTGTACAGCCACTGTACCCATTTCTGGAGATGTTTCCCCAGTA	1080
QY	1023	TCCACTTGGCTCCCCGAGCCCCAGTTTGGAGGAGGGGTGCCGCAACACAGAGTCCCT	1082
Db	1081		
		TCCACTTGGCTCCCCGAGCCCCAGTTTGGAGGAGGGGTGCCGCAACACAGAGTCCCT	1140
QY	1083	CTGGACTTGACACGAGGAGCCAGTTGGAACCCGGGAGCAGAGCCAGGTGGGCCACCGT	1142
Db	1141		
		CTGGACTTGACACGAGGAGCCAGTTGGAACCCGGGAGCAGAGCCAGGTGGGCCACCGT	1200
QY	1143	ACCAATGGCATTTATGTCACCGCGGGTCTATGACTATCTACTGGCAACATCTACATCTAC	1202
Db	1201		
		ACCAATGGCATTTATGTCACCGCGGGTCTATGACTATCTACTGGCAACATCTACATCTAC	1260
QY	1203	AATGGACCACTACTGGGGGACACCGGGTCTGTGAGACCTCCAGCTACCCCGGAACCT	1262
Db	1261		
		AATGGACCACTACTGGGGGACACCGGGTCTGTGAGACCTCCAGCTACCCCGGAACCT	1320
QY	1263	CCATACCCCATTCGCCGAAGAGGGGACCTGCGCCCTCCCGGGTCTCTTACACCCACACAG	1322
Db	1321		
		CCATACCCCATTCGCCGAAGAGGGGACCTGCGCCCTCCCGGGTCTCTTACACCCACACAG	1380
QY	1323	GAAGATGGCAAGGCTTTGGCACTAGCGGAGACAGAGCACTGTGGTGGCCACACCCCTTAAC	1382

Db	1381	GAAGATGGCAAGGCTTGGCACTTAGCGGAGACAGAGCACTGTGGTGCCACACCCCTCTTAAC	1440
Qy	1383	AGGGGCCCAAGGAACCAATTTATCACCCATGACTGACTGAGTCTGAGAAAAGGCAGAAGA	1442
Db	1441	AGGGGCCCAAGGAACCAATTTATCACCCATGACTGACGGAGTCTGAGAAAAGGCAGAAGA	1500
Qy	1443	AGGGGGGCACAAGGGCAACCTTCTCCCTTGAGGCTGCCCTGCCAGCTGGGATTCACAGGG	1502
Db	1501	AGGGGGGCACAAGGGCACTTTCTCCCTTGAGGCTGCCCTGCCAGCTGGGATTCACAGGG	1560
Qy	1503	GCCTGAGTAGGGCCCGGGGAAGCAGAGCCCTTAAGGGATTAAAGGCTCAGACACCTCTGAGA	1562
Db	1561	GCCTGAGTAGGGCCCGGGGAAGCAGAGCCCTTAAGGGATTAAAGGCTCAGACACCTCTGAGA	1620
Qy	1563	GCAGGTGGGCACCTGGCTGGGTGAGGTGCCCTCCACAGGACTCTCCCTACTCGCTGAGCAA	1622
Db	1621	GCAGGTGGGCACCTGGCTGGGTGAGGTGCCCTCCACAGGACTCTCCCTACTCGCTGAGCAA	1680
Qy	1623	ACCTGAGGCCCTCCCGGAGACCCACCCCTCGGGGCTGCTCAGCCTCAGGCACCGGAC	1682
Db	1681	ACCTGAGGCCCTCCCGGAGACCCACCCCTCGGGGCTGCTCAGCCTCAGGCACCGGAC	1740
Qy	1683	AGGGCACAATGATACCACTGCTGCCACACTACAGCAGCGCGGACACGGACCGGCACCGAG	1742
Db	1741	AGGGCACAATGATACCACTGCTGCCACACTACGGCAGCGCGGACACGGGACCGGCACCGAG	1800
Qy	1743	GGAGCGGCACACGGTCACTGCAAGGAGCGTCAAGGGGCCCTCTAAAGGATTTCGTGGTGCC	1802
Db	1801	GGAGCGGCACACGGTCACTGCAAGGAGCGTCAAGGGGCCCTCTAAAGGATTTCGTGGTGCC	1860
Qy	1803	TCATCCCCAAGCTTCAGAGACCCCTTTGGGTTCCACACTTCACGTGGGACTGAGGTAGAAC	1862
Db	1861	TCATCCCCAAGCTTCAGAGACCCCTTTGGGTTCCACACTTCACGTGGGACTGAGGTAGAAC	1920
Qy	1863	CTGCATGAAGTGAATATAGGGAGAGCGCTCTCCCTCCCTCCTCTAGAGGAGAGGAA	1922
Db	1921	CTGCATGAAGTGAATATAGGGAGAGCGCTCTCCCTCCCTCCTCTAGAGGAGAGGAA	1980
Qy	1923	AGGGAGTCATTAAACAACCTAGGGGTTGGGTAGGATTCCTAGTATGCGGAAGAGTTTGG	1982
Db	1981	AGGGAGTCATTAAACAACCTAGGGGTTGGGTAGGATTCCTAGTATGCGGAAGAGTTTGG	2040
RESULT 5			
ABQ55003			
ID	ABQ55003 standard; cDNA; 1594 BP.		
XX			
AC	ABQ55003;		
XX			
DT	22-AUG-2002 (first entry)		
XX			
DE	Human ovarian antigen HSRBJ44 cDNA, SEQ ID NO:883.		
XX			
KW	Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;		
KW	ovarian cancer; breast cancer; tumour; reproductive system disorder;		
KW	infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;		
KW	PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;		
KW	inflammatory condition; immune disorder; blood disorder;		
KW	cardiovascular disorder; respiratory disorder; neurological disorder;		
KW	gastrointestinal disorder; urinary system disorder; drug screening;		
KW	gene therapy; chromosome mapping; forensic analysis;		
KW	antibody preparation; cytostatic; immunomodulatory; neuroprotective;		
KW	antiinflammatory; gynaecological; reproductive; chromosome 12p13;		
KW	gene; ss.		
OS	Homo sapiens.		
XX			
PN	WO200200677-A1.		
XX			
PD	03-JAN-2002.		
XX			
PF	07-JUN-2001; 2001WO-US1569.		

	07-JUN-2000; 2000US-209467P.
(HUMA-) HUMAN GENOME SCI INC.	
Birse CE, Rosen CA;	
WPI; 2002-147878/19.	
P-PSDB; ABP41926.	
Isolated nucleic acid molecules encoding novel ovarian polypeptides, useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian cancer), immune disorders, cardiovascular disorders and neurological diseases	
Claim 1; SEQ ID No 883; 2922pp; English.	
The invention relates to 2175 novel human ovarian antigens (ABP41054- ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also encompasses polypeptides 90% identical and polynucleotides 95% identical to the sequences of the invention. The invention additionally relates to recombinant vectors and host cells comprising human ovarian antigen polynucleotides, antibodies against human ovarian antigens, and the use of ovarian antigen polynucleotides and polypeptides in diagnosing, treating, prognosing or preventing various ovary and/or breast-related disorders. Such conditions include ovarian cancer and breast cancer, and metastatic tumours of ovarian or breast origin, reproductive system disorders (e.g., infertility, disorders of pregnancy, anovulation, polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and vaginitis), immune disorders (e.g., congenital and acquired immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus), blood-related disorders (e.g., anaemia), cardiovascular disorders, respiratory disorders, neurological disorders, gastrointestinal disorders and urinary system disorders. Ovarian antigen polypeptides and polynucleotides may also be used in screening for compounds which modulate ovarian antigen expression or activity. The polynucleotides may further be used for gene therapy, chromosome mapping, in the identification of individuals and in forensic analysis, and the polypeptides may be used as food additives or to prepare antibodies useful in disease diagnosis, drug targeting and phenotyping. The present sequence represents cDNA encoding a human ovarian antigen of the invention.	
Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences..	
Sequence 1594 BP; 391 A; 488 C; 434 G; 280 T; 1 other;	
Query Match 59.2%; Score 1173.4; DB 24; Length 1594; Best Local Similarity 92.2%; Pred. No. 4.7e-261; Matches 1292; Conservative 0; Mismatches 2; Indels 108; Gaps 1	
Yq 689 AGATGAAGTTGGGAAAGGTAAACAACACTGCCTGCCCTGCAAGCGAGCCGACTTCCAGAA 748 	
Dd 91 AGATGAAGTTGGGAAAGGTAAACAACACTGCCTGCCCTGCAAGCGAGCCGACTTCCAGAA 150 	
Yq 749 TAOCCTCCTCCCAGGCCCCCGTGTCAGGCCACCACACAGGTGTGAGAACCAGAGTCTGGT 808 	
Dd 151 TACCTCCTCCCAGGCCCGCGTGCAGGCCCACACACAGGTGTGAGAACCAGAGTCTGGT 210 	
Yq 809 GGAGGAGCTCCAGGCACTGCCAGTCCACACACACTGCAAAATCCATTAGAGCCACT 868 	
Dd 211 GGAGGAGCTCCAGGCACTGCCAGTCCACACACACTGCAAAATCCATTAGAGCCACT 270 	
Yq 869 GCCCCCCAGAGATGTCA----- 884 	
Dd 271 GCCCCCAGAGATGTCAAGAACCATGCTGATGTGCCCGTCTGCTGCCACTGGCCTTCTT 330 	
Yq 885 ----- 884	
Dd 331 TCTGCTCCTTGCCACCGCTCTTCTCCTGCATCTGGAAGAGCCACCCTTCTCTCTGCAGAA 390 	

QY 1961 TAGGTATGGGAGAGATTGG 1982
|||||
Db 1471 TAGGTATGGGAGAGATTGG 1492

RESULT 6
ABK54540
ID ABK54540 standard; cDNA; 527 BP.

XX ABK54540;

XX 18-JUN-2002 (first entry)

XX Human colon cancer-associated cDNA, SEQ ID No 10.

XX Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.

XX Homo sapiens.

XX WO200212280-A2.

PN 14-FEB-2002.

XX 30-JUL-2001; 2001WO-US23826.

XX 03-AUG-2000; 2000US-223265P.

PR 02-OCT-2000; 2000US-237406P.

PR 20-MAR-2001; 2001US-277495P.

PR 03-JUL-2001; 2001US-302702P.

XX (CORI-) CORIXA CORP.

PA Pyle RA, Xu J, Secrist H;

XX WPI; 2002-257462/30.

XX Novel polynucleotide encoding colon tumour polypeptides, useful as
vaccines for treating colon cancers

PS Claim 1; Page 147; 425pp; English.

XX The invention relates to isolated polynucleotides (I) encoding colon
tumour polypeptides (II). (I) is useful for stimulating an immune
response in a patient and treating colon cancer in a patient.
XX Oligonucleotides derived from (I) are useful for determining the presence
of cancer in a patient. (I) and (II) are useful in pharmaceutical
compositions, e.g. vaccines, and other compositions for the diagnosis
and treatment of colon cancer. A composition comprising a first component
selected from physiologically acceptable carriers and immunostimulants,
and an antigen-presenting cell expressing (II) is useful for inhibiting
development of cancer in a patient. (I) is useful in the design and
preparation of ribozyme molecules for inhibiting expression of tumour
polypeptides and (I). ABK54531-ABK55464 represent human colon cancer cDNA
sequences of the invention.

XX SQ Sequence 527 BP; 86 A; 190 C; 158 G; 88 T; 5 other;

Query Match 22.4%; Score 444.6; DB 24; Length 527;
Best Local Similarity 88.1%; Pred. No. 4.9e-93;
Matches 518; Conservative 0; Mismatches 9; Indels 61; Gaps 1;

QY 47 GGGAGCCCTGGAGCCCGGCTGGCCCTCCCGGCCCTGGGGTGCACATCGGCCCTGAGT 106

||||| 1 GGGAGCNCCTGGAGCCCGGCTGGCCCTCCCGGCCCTGGGGTGCACATCGGCCCTGAGT 60

QY 107 CCGGTCCAGGCTGTGGGCTGGGAGCCGCCGCCACCGTGGCCAGGACGTGGGGCTC 166

||||| 61 CCGGTCCAGGCTGTGGGCTGGGAGCCGCCGCCACCGTGGCCAGGACGTGGGGCTC 89

QY 167 CTGCTTCCTCCAGGCCGCCACGTTGCTGGCGCCCTGGCGAGTGGCCGCATGCTCCT 226

||||| 90 -----GCCGCTGGCGGAGTGGCCGCATGCTCCT 119

QY 227 GCCTTGGGCCACCTCTGCCCGCCCGGCTGGCGCTGGGGGCGCTCTGGTGTGGGCCCTCTTCGG 286
|||||
Db 120 GCCTTGGGCCACCTCTGCCCGCCCGGCTGGCGCTGGGGGCGCTCTGGTGTGGGCCCTCTTCGG 179
QY 287 GCTCCTGGCAGCATCGCAGCCCGGCGGTGCTCCATATGCGTGGGAGACGACCTG 346
|||||
Db 180 GCTCCTGGCAGCATCGCAGCCCGGCGGTGCTCCATATGCGTGGGAGACGACCTG 239
QY 347 CAGGACCCAGGAAAGAAATATAGCCCGCCAGCCGATCTGCTGCTCCCGCTGCC 406
|||||
Db 240 CAGGACCCAGGAAAGAAATATAGCCCGCCAGCCGATCTGCTGCTCCCGCTGCC 299
QY 407 GCCAGGCACCTATGCTCAGCTAAATGTAGCCGATCCGGGACACAGTTTGTGCCACATG 466
|||||
Db 300 GCCAGGCACCTATGCTCAGCTAAATGTAGCCGATCCGGGACACAGTTTGTGCCACATG 359
QY 467 TGCCGAGAAATTCCTACACGAGCAGTGAAGTCTGACCATCTGCCAGCTGTGCCGCC 526
|||||
Db 360 TGCCGAGAAATTCCTACACGAGCAGTGAAGTCTGACCATCTGCCAGCTGTGCCGCC 419
QY 527 CTGTGACCCAGTGTGGCGCTCGAGGAGATTGCTGTGCTGCTGGGCCCTCGAGTG 634
|||||
Db 420 CTGTGACCCAGTGTGGCGCTCGAGGAGATTGCTGTGCTGCTGGGCCCTCGAGTG 527
QY 587 GTGCCGCTGCCAGCCGGGAAATGTTCTGTGCTGCTGGGCCCTCGAGTG 634
|||||
Db 480 GTGCCGCTGCCAGCCGGGAAATGTTCTGTGCTGCTGGGCCCTCGAGTG 527

RESULT 7

ABL81903

ID ABL81903 standard; cDNA; 481 BP.

XX ABL81903;

XX 17-MAY-2002 (first entry)

XX Human ovarian cancer related cDNA clone SEQ ID NO:4881.

XX Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.

XX Homo sapiens.

XX WO200192581-A2.

XX 06-DEC-2001.

XX 29-MAY-2001; 2001WO-US17756.

XX 26-MAY-2000; 2000US-207484P.

XX (CORI-) CORIXA CORP.

XX Algate PA, Harlocker SL, Jones R;

XX WPI; 2002-122075/16.

XX Composition for therapy and diagnosis of ovarian cancer comprising
polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
polypeptide, antibody specific to polypeptide or T cell expressing
polypeptide

PS Claim 1; SEQ ID 4881; 489pp; English.

XX The present invention describes a composition (I) comprising: carriers
and immunostimulants; and a polypeptide (II) of a ovarian tumour .
XX polypeptide encoded by a polynucleotide (III) having a cDNA sequence
(S1) from the 10912 nucleotide sequences as given in ABU77023 to
CC ABL87934, (III) encoding (II) having a sequence (S2), a T cell
population of (II), or antigen presenting cells that express (II).
XX (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to
CC (S1) can be used for detecting ovarian cancer in a patient's biological
sample preferably serum or ovarian tissue. The method comprises

CC contacting a biological sample from a patient with (IV), detecting the
 CC amount of polynucleotide hybridising to (IV) and comparing the amount to
 CC a predetermined cutoff value and thereby detecting ovarian cancer in the
 CC patient, where the amount of polynucleotide hybridising to (IV) is
 CC detected preferably by polymerase chain reaction (PCR). (I) comprising
 CC (III) and/or (II) is useful for stimulating and/or expanding T cells
 CC specific for an ovarian tumour protein comprising contacting T cells
 CC with (III) or (II). (III) is useful in design and preparation of
 CC ribozyme molecules for inhibiting expression of the tumour polypeptides
 CC and proteins in tumour cells; and to isolate a full length gene from a
 CC suitable library e.g., a tumour cDNA library using well known
 CC techniques.

XX Sequence 481 BP; 113 A; 161 C; 123 G; 84 T; 0 other;

Query Match 21.9%; Score 433.6; DB 24; Length 481;
 Best Local Similarity 96.5%; Pred. No. 1.6e-90;
 Matches 465; Conservative 0; Mismatches 14; Indels 3; Gaps 2;

QY 390 TCGTCTCCGCTGCGCGCCAGGACCTATGCTCAGCTAAATGTAGCGCATCCGGGAC 449

Db 1 TCGTCTCCGCTGCGCGCCAGGACCTATGCTCAGCTAAATGTAGCGCATCCGGGAC 59

QY 450 ACAGTTTGTGCGACATGTCGCGAGATTCCTACAAGGACCTGGAACCTGACCATC 509

Db 60 ACAGTTTGTGCGACATGTCGCGAGATTCCTACAAGGACCTGGAACCTGACCATC 119

QY 510 TCCAGCTGTGCGCCCTGTGACCCAGTGTGCGGCTCGAGGAGATTGCCCTTCGACA 569

Db 120 TCCAGCTGTGCGCCCTGTGACCCAGTGTGCGGCTCGAGGAGATTGCCCTTCGACA 179

QY 570 AGCAACGGAGAACCCAGTGGCGTGCAGCGGGGAATGTTCTGTGCTGGGCCCTC 629

Db 180 AGCAACGGAGAACCCAGTGGCGTGCAGCGGGGAATGTTCTGTGCTGGGCCCTC 239

QY 630 GAGTGTACACACTGCGAGTACTTTCTGACTGCCCGCTGGCACTGAAGCCGAGCTCAA 689

Db 240 GAGTGTACACACTGCGAGTACTTTCTGACTGCCCGCTGGCACTGAAGCCGAGCTCAA 299

QY 690 GATGAAGTTGGAAGGTTAAACCACTGCTGCTCCCTGCAAGGCGGACATTT--CCAGA 747

Db 300 GATGAAGTTGGAAGGTTAAACCACTGCTGCTCCCTGCAAGGCGGACATTTCCAGAA 359

QY 748 ATACCTCTCTCCCGAGCGCGCTGCCAGCCGCCACACAGGTGTGAGAACCAAGTCTGG 807

Db 360 GTACCTCTCTCCCGAGCGCGCTGCCAGCCGCCACACAGGTGTGAGAACCAAGTCTGG 419

QY 808 TGGAGGCGAGCTCCAGGCACTGCGGCTGCGGACACCACTGCAAAAATCCATTAGAGCCAC 867

Db 420 TGGAGGCGAGCTCCAGGCACTGCGGCTGCGGACACCACTGCAAAAATCCATTAGAGCCAC 479

QY 868 TG 869

Db 480 TG 481

RESULT 8

ABL82725

ID ABL82725 standard; cDNA; 529 BP.

XX

AC ABL82725;

XX

DT 17-MAY-2002 (first entry)

XX Human ovarian cancer related cDNA clone SEQ ID NO:5703.

XX Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.

XX Homo sapiens.

XX WO200192581-A2..

XX 06-DEC-2001.

PD

XX

PF 29-MAY-2001; 2001WO-US17756.

PR

XX 26-MAY-2000; 2000US-207484P.

XX

PA (CORI-) CORIXA CORP.

XX

PI Algate PA, Harlocker SL, Jones R;

DR

XX WPI; 2002-122075/16.

XX

Composition for therapy and diagnosis of ovarian cancer comprising

polypeptide of a ovarian tumor polypeptide, polynucleotide encoding

polypeptide, antibody specific to polypeptide or T cell expressing

polypeptide

PS Claim 1; SEQ ID 5703; 489pp; English.

XX

The present invention describes a composition (I) comprising: carriers and immunostimulants; and a polypeptide (II) of a ovarian tumour polypeptide encoded by a polynucleotide (III) having a cDNA sequence (SI) from the 10912 nucleotide sequences as given in ABL77023 to ABL87934, (III) encoding (II) having a sequence (S2), a T cell population of (II), or antigen presenting cells that express (II). (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to (SI) can be used for detecting ovarian cancer in a patient's biological sample preferably serum or ovarian tissue. The method comprises contacting a biological sample from a patient with (IV), detecting the amount of polynucleotide hybridising to (IV) and comparing the amount to a predetermined cutoff value and thereby detecting ovarian cancer in the patient, where the amount of polynucleotide hybridising to (IV) is detected preferably by polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is useful for stimulating and/or expanding T cells specific for an ovarian tumour protein comprising contacting T cells with (III) or (II). (III) is useful in design and preparation of ribozyme molecules for inhibiting expression of the tumour polypeptides and proteins in tumour cells; and to isolate a full length gene from a suitable library e.g., a tumour cDNA library using well known techniques.

XX Sequence 529 BP; 84 A; 185 C; 149 G; 109 T; 2 other;

Query Match 19.5%; Score 386.4; DB 24; Length 529;

Best Local Similarity 88.8%; Pred. No. 1.3e-79;

Matches 474; Conservative 0; Mismatches 53; Indels 7; Gaps 5;

QY 76 CCGGCGCTGGGTCACATCGGCCTGAGTCCCTCCAGGCTCGGGCTGGGCGAGCC 135

Db 1 CCGGCGCTGGGTCACATCGGCCTGAGTCCCTCCAGGCTCGGGCTGGGCGAGCC 60

QY 136 GCGGCGACCGTGGCCAGGAGTGGGGCTCTGCTCCAGGCGCCCGCCAGCTGCT 195

Db 61 G-CGCCANCGTGGCCAGGAGTGGGGCTCTGCTCCAGGCGCCCGCCAGCTGCT 119

QY 196 GCGGCGCTGGCGAGTGGCCCGCCATG--CTCCTGCTGGGCGACCTGTGCCCGGCT 253

Db 120 CCTCTCGGAGCTGGCGGCTCATGCTCTTGGGGGACCTGTGCCCGGCT 179

QY 254 GCGCTGGGGGCTCTGCTGGTGGGCTCTCGGGCTCTCGGAGCATCGAGCCCGGCG 313

Db 180 GCGCTGGGCG--CTCTGGTGGGCTCTCGGGCTCTCGGAGCATCGAGCCCGGCG-- 236

QY 314 GGTGCTCCATATGCTCGGAGAACACAGCTGCGAGGACCGAGGAGGAAATATGTA 373

Db 237 CGGTGCTCCATATGCTCGGAGAACACAGCTGCGAGGACCGAGGAGGAAATATGTA 296

QY 374 GCGGCGACCGCATCTGCTCTCCGCTGCGCGGAGGACCTATGCTCAGCTAAATG 433

Db 297 GCGGCGACCGCATCTCTCTCCGCTGCGCGGAGGACCTATGCTCAGCTAAATG 355

QY 434 TAGCGCATCGGGGACACAGTTTGTGCCACATGTGCCAGAAATTCCTACAAGGACCTG 493

Db 356 TAGCGCATCGGGGACACAGTTTGTGCCACATGTGCCAGAAATTCCTACAAGGACCTG 415

QY 494 GAACTACTGACCATCTGCAGCTGTGCGCCCTCTGTGACCCAGTGTATGGCCCTCGAGGA 553
|||||
Db 416 GAACTACTGACCATCTGCAGCTGTGCGCCCTCTGTGACCCAGTGTATGGCCCTCGAGGA 475
QY 554 GATTGCCCTGTGACAGCAAGCAAGCGAGACCCAGTGTGCGCTGCCAGCGGGAAT 607
|||||
Db 476 GATTGCCCTGTGACAGCAAGCAAGCGAGACCCAGTGTGCGCTGCCAGCGGGAAT 529

RESULT 9
ABL82484
ID ABL82484 standard; cDNA; 516 BP.
XX
AC
ABL82484;
17-MAY-2002 (first entry)
DE Human ovarian cancer related cDNA clone SEQ ID NO:5462.
XX
DE Human ovarian cancer; ovarian tumour; cytostatic; gene; ss.
KW Homo sapiens.
OS
XX
XX
PN WO200192581-A2.
XX
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US17756.
XX
PR 26-MAY-2000; 2000US-207484P.
XX
PA (CORI-) CORIXA CORP.
PI Algate PA, Harlocker SL, Jones R;
XX
XX WPI; 2002-122075/16.
XX
XX
XX Composition for therapy and diagnosis of ovarian cancer comprising polypeptide of a ovarian tumor polypeptide, polynucleotide encoding polypeptide, antibody specific to polypeptide or T cell expressing polypeptide -
PS Claim 1; SEQ ID 5462; 489pp; English.
XX
XX The present invention describes a composition (I) comprising: carriers and immunostimulants; and a polypeptide (II) of a ovarian tumour polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1) from the 10912 nucleotide sequences as given in ABL77023 to ABL87934, (III) encoding (II) having a sequence (S2), a T cell population of (II), or antigen presenting cells that express (II). (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to (S1) can be used for detecting ovarian cancer in a patient's biological sample preferably serum or ovarian tissue. The method comprises contacting a biological sample from a patient with (IV), detecting the amount of polynucleotide hybridising to (IV) and comparing the amount to a predetermined cutoff value and thereby detecting ovarian cancer in the patient, where the amount of polynucleotide hybridising to (IV) is detected preferably by polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is useful for stimulating and/or expanding T cells specific for an ovarian tumour protein comprising contacting T cells with (III) or (II). (III) is useful in design and preparation of ribozyme molecules for inhibiting expression of the tumour polypeptides and proteins in tumour cells; and to isolate a full length gene from a suitable library e.g., a tumour cDNA library using well known techniques.

XX
XX Sequence 516 BP; 83 A; 187 C; 153 G; 93 T; 0 other;
Query Match 19.3%; Score 382; DB 24; Length 516;
Best Local Similarity 90.9%; Pred. NO. 1.3e-78;
Matches 476; Conservative 0; Mismatches 40; Indels 10; Gaps 6;

QY 95 TCGGCCCTGAGTCCCGTCCAGGCTCTGGGCTCGGGCAGCGCGCCACCGCTGCCAGG 154
|||||
Db 1 TCGGCCCTGAGTCCCGTCCAGGCTCT -GGCTCGGGCAGCG -CGCCACCGCTGCCAGG 58
QY 155 ACGTCGGGCTCTGCTTCTCCAGGCCCCACGCTTGCTGGCCGCTGCCAGTGGC 214
|||||
Db 59 ACGTCGGGCTCTGCTTCTCCAGGCCCCACGCTTGCTGGCCGCTGCCAGGCTGGC 118
|||||
QY 215 CGCATCTCTGCTTGGGCGCACCTCTGCCCGGCTGGCTTGGGGGCTCTGTGGTCT 274
|||||
Db 119 GCGATGCTCTGCTTGGGCGCACCTCTGCCCGGCTGGCTTGGGGGCTCTGTGGTCT 175
|||||
QY 275 GGGCTCTTTCGGGCTCTGTCAGCATCGAGCCCCAGCGGCTGCTCATATGCGTCGA 334
|||||
Db 176 GGGCTCTTTCGGGCTCTGTCAGCATCGAGCCCCAGCGGCTGCTCATATGCGTCGA 235
|||||
QY 335 GAACGAGACCTGCGAGGACCCAGGAAAGAAATGACTATGAGCCCCAGCGCATCTGCTG 394
|||||
Db 236 GAACGAGACCTGCGAGGACCCAGGAAAGAAATGACTATGAGCCCCAGCGCATCTGCTG 293
|||||
QY 395 CTCCCGCTGCGCGGCGCACCTATGCTCAGCTAAATGTAGCGCATCGGGGACACAGT 454
|||||
Db 294 CGCCCGCTGCGCGGCGCACCTATGCTCAGCTAAATGTAGCGCATCGGGGACACAGT 350
|||||
QY 455 TTGTGCCACATGTGCCGAGAAATTCCTACACGAGCACTGGAACCTACCTGACCACTGCCA 514
|||||
Db 351 TTGTGCCACATGTGCCGAGAAATTCCTACACGAGCACTGGAACCTACCTGACCACTGCCA 410
|||||
QY 515 GCTGTGCCGCGCTGTGACCCAGTGTGGGCTCGAGGAGATGCCCCCTGCACAGCAA 574
|||||
Db 411 GCTGTGCCGCGCTGTGACCCAGTGTGGGCTCGAGGAGATGCCCCCTGCACAGCAA 470
|||||
QY 575 ACGGAAGACCCAGTGGCGCTGCCAGCGGGAATGTTCTGTGCTGCC 620
|||||
Db 471 ACGGAAGACCCAGTGGCGCTGCCAGCGGGAATGTTCTGTGCTGCC 516
|||||

RESULT 10
ABL82623
ID ABL82623 standard; cDNA; 281 BP.
XX
AC ABL82623;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human ovarian cancer related cDNA clone SEQ ID NO:5601.
XX
KW Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200192581-A2.
XX
PD 06-DEC-2001.
XX
XX 29-MAY-2001; 2001WO-US17756.
PF
XX 26-MAY-2000; 2000US-207484P.
PR
XX (CORI-) CORIXA CORP.
PA Algate PA, Harlocker SL, Jones R;
XX
PI WPI; 2002-122075/16.
XX
XX Composition for therapy and diagnosis of ovarian cancer comprising polypeptide of a ovarian tumor polypeptide, polynucleotide encoding polypeptide, antibody specific to polypeptide or T cell expressing polypeptide -
PS Claim 1; SEQ ID 5601; 489pp; English.
XX
XX The present invention describes a composition (I) comprising: carriers

CC and immunostimulants; and a polypeptide (II) of an ovarian tumour
 CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence
 CC (SI) from the 10912 nucleotide sequences as given in ABL77023 to
 CC ABL87934, (III) encoding (II) having a sequence (S2), a T cell
 CC population of (II), or antigen presenting cells that express (II).
 CC (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to
 CC (SI) can be used for detecting ovarian cancer in a patient's biological
 CC sample preferably serum or ovarian tissue. The method comprises
 CC contacting a biological sample from a patient with (IV), detecting the
 CC amount of polynucleotide hybridising to (IV) and comparing the amount to
 CC a predetermined cutoff value and thereby detecting ovarian cancer in the
 CC patient, where the amount of polynucleotide hybridising to (IV) is
 CC detected preferably by polymerase chain reaction (PCR). (I) comprising
 CC specific for an ovarian tumour protein comprising contacting T cells
 CC with (III) or (II). (III) is useful in design and preparation of
 CC ribozyme molecules for inhibiting expression of the tumour polypeptides
 CC and proteins in tumour cells; and to isolate a full length gene from a
 CC suitable library e.g., a tumour cDNA library using well known
 CC techniques.
 XX SQ Sequence 281 BP; 58 A; 94 C; 77 G; 52 T; 0 other;
 Query Match 12.8%; Score 253.8; DB 24; Length 281;
 Best Local Similarity 97.1%; Pred. No. 4.4e-49;
 Matches 269; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
 QY 996 CCATTCTCGAGATGTTCCCACTATCCACTGGGCTCCCGCAGCCC-CAGTTTGGGA 1054
 DB 5 CCATTCTCGAGATGTTCCCACTATCCACTGGGCTCCCGCAGCCCAGTTTGGGA 64
 QY 1055 GCGAGGGTGCCGCAACAGCAGAGTCTCTGGACCTGACCGAGGAGCGCGAGTTGGAACC 1114
 DB 65 GCGAGGGTGCCGCAACAGCAGAGTCTCTGGACCTGACCGAGGAGCGCGAGTTGGAACC 124
 QY 1115 CGGGAGCAGACCGAGTGGCCCGACGGTACCAATGGCATTCATGTCACCGCGGGTCTAT 1174
 DB 125 CGGGAGCAGACTCAGTGGCCCGACGGTACCAATGGCATTCATGTCACCGCGGGTCTAT 184
 QY 1175 GACTATCACTGGCAACACTTACATCTACAATGGACAGTACTGGGGGACACCGGGTCC 1234
 DB 185 GACTATCACTGGCAACACTTACATCTACAATGGACAGTACTGGGGGACACCGGGTCA 244
 QY 1235 TGGAGACTCCCGAGTACCCCGCAACCTCCATACCCC 1271
 DB 245 TGGAGACTCCCGAGTACCCCGCAACCTCCATTCGCC 281

RESULT 11
 ABL82688/c
 ID ABL82688 standard; cDNA; 289 BP.
 XX AC ABL82688;
 XX DT 17-MAY-2002 (first entry)
 XX DE Human ovarian cancer related cDNA clone SEQ ID NO:5666.
 XX KW Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
 XX OS Homo sapiens.
 XX PN WO200192581-A2.
 XX PD 06-DEC-2001.
 XX PF 29-MAY-2001; 2001WO-US17756.
 XX PR 26-MAY-2000; 2000US-207484P.
 XX PA (CORI-) CORIXA CORP.
 XX PI Algate PA, Harlocker SL, Jones R;

XX WPI; 2002-122075/16.
 XX Composition for therapy and diagnosis of ovarian cancer comprising
 XX polypeptide of an ovarian tumor polypeptide, polynucleotide encoding
 XX polypeptide, antibody specific to polypeptide or T cell expressing
 XX polypeptide
 XX Claim 1; SEQ ID 5666; 489pp; English.
 XX The present invention describes a composition (I) comprising: carriers
 XX and immunostimulants; and a polypeptide (II) of an ovarian tumour
 XX polypeptide encoded by a polynucleotide (III) having a cDNA sequence
 XX (SI) from the 10912 nucleotide sequences as given in ABL77023 to
 XX ABL87934, (III) encoding (II) having a sequence (S2), a T cell
 XX population of (II), or antigen presenting cells that express (II).
 XX (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to
 XX (SI) can be used for detecting ovarian cancer in a patient's biological
 XX sample preferably serum or ovarian tissue. The method comprises
 XX contacting a biological sample from a patient with (IV), detecting the
 XX amount of polynucleotide hybridising to (IV) and comparing the amount to
 XX a predetermined cutoff value and thereby detecting ovarian cancer in the
 XX patient, where the amount of polynucleotide hybridising to (IV) is
 XX detected preferably by polymerase chain reaction (PCR). (I) comprising
 XX specific for an ovarian tumour protein comprising contacting T cells
 XX with (III) or (II). (III) is useful in design and preparation of
 XX ribozyme molecules for inhibiting expression of the tumour polypeptides
 XX and proteins in tumour cells; and to isolate a full length gene from a
 XX suitable library e.g., a tumour cDNA library using well known
 XX techniques.
 XX SQ Sequence 289 BP; 49 A; 75 C; 93 G; 72 T; 0 other;
 Query Match 12.7%; Score 251.2; DB 24; Length 289;
 Best Local Similarity 98.5%; Pred. No. 1.7e-48;
 Matches 264; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 QY 1007 AGATGTTTCCCGACTATCCACTGGGCTCCCGCAGCCCGAGTTTGGAGGAGGGTGCC 1066
 DB 289 AGATGTTTCCCGACTATCCACTGGGCTCCCGCAGCCCGAGTTTGGAGGAGGGTGCC 230
 QY 1067 GCAACAGCAGA-GTCTCTGGACCTGACCGAGGAGCGCGAGTTGGAACCGGGAGCAGA 1125
 DB 229 GCAACAGCAGATGTCTCTGGACCTGACCGAGGAGCGCGAGTTGGAACCGGGAGCAGA 170
 QY 1126 GCCAGTGGCCCGACGGTACCAATGGCATTCATGTCACCGCGGGTCTATGACTATCACTG 1185
 DB 169 GCCAGTGGCCCGACGGTACCAATGGCATTCATGTCACCGCGGGTCTATGACTATCACTG 110
 QY 1186 GCAACATCTACATCTACAATGGACAGTACTGGGGGACACCGCGGTCTCTGGAGACCTCC 1245
 DB 109 GCAACATCTACATCTACAATGGACAGTACTGGGGGACACCGCGGTCTCTGGAGACCTCC 50
 QY 1246 CAGTACCCCGGAACCTCCATACCCCAT 1273
 DB 49 CAGTACCCCGGAACCTCCATTCGCCCT 22

RESULT 12
 AAQ45225
 ID AAQ45225 standard; cDNA; 1557 BP.
 XX AC AAQ45225;
 XX DT 07-OCT-1994 (first entry)
 XX DE Sequence encoding a recombinant human tumour necrosis factor receptor
 XX TNFR/fc fusion protein.
 XX KW Tumour necrosis factor receptor; chimeric antibody molecule;
 XX immunoglobulin; ss.

Db	335	ACACCCAGCTCTCGAAACTGGGGTCCCGAGTGTCTGAGCTGTGGCTCCCG--CTGTAGCT	391			
QY	541	TGGGCTCTCGAGGAGATGGCCCCCTGCACAAAGCAAAACGGAAGACCCAGTCCGGCTGCCAGC	600			
Db	392	GTGACCAGGTGGAAACTCAAGCTTGACCTCGGACACAGAACCGATCTGCACCTGCAGGC	451			
QY	601	CGGGAATGTTCTGTGC	616			
Db	452	CCGGCTGGTACTGCGC	467			
RESULT 13						
AAC90446						
ID	AAC90446 standard; cDNA; 1557 BP.					
XX	AAC90446;					
AC						
XX	19-MAR-2001 (first entry)					
DT						
XX	TNFR:Fc fusion protein coding sequence.					
DE						
XX						
XX	Antiinflammatory; antiarthritic; gene therapy; inflammatory disorder;					
KW	Tumour Necrosis Factor receptor; TNFR; IgG1 Fc domain; asthma;					
KW	rheumatoid arthritis; Crohn's disease; congestive heart failure; ss.					
KW						
XX	Rattus sp.					
OS						
XX	W0200073481-A1.					
PN						
XX	07-DEC-2000.					
PD						
XX						
XX	26-MAY-2000; 2000WO-US14586.					
PF						
XX						
XX	28-MAY-1999; 99US-0150688.					
PR						
XX						
XX	(TARG-) TARGETED GENETICS CORP.					
PA						
XX						
PI	Burstein H, Stepan AM;					
XX						
XX	WPI: 2001-061552/07.					
DR						
DR	P-PSDB: AAB50080.					
PT	Recombinant adeno-associated virus vector, useful for palliating tumor					
PT	necrosis factor associated disorder, comprises polynucleotide which					
PT	encodes fusion polypeptide comprising tumor necrosis factor receptor					
XX						
XX	Example 1; Fig 2; 90pp; English.					
PS						
XX						
CC	The present invention relates to a recombinant adeno-associated virus					
CC	(rAAV) vector comprising a polynucleotide (the present sequence) which					
CC	encodes a fusion protein comprising an extracellular domain of Tumour					
CC	Necrosis Factor receptor (TNFR) and a constant domain (Fc) of IgG1					
CC	molecule. The vector of the present invention is useful for reducing T					
CC	levels, for reducing an inflammatory response which occurs in a					
CC	connective tissue and for palliating a TNF-associated disorder such as					
CC	inflammatory disorders e.g. rheumatoid arthritis, Crohn's disease, as					
CC	and congestive heart failure.					
XX						
SQ	Sequence 1557 BP; 345 A; 526 C; 445 G; 241 T; 0 other;					
Query Match 3.6%; Score 72; DB 22; Length 1557;						
Best Local Similarity 51.3%; Pred. No. 5.7e-07;						
Matches 193; Conservative 0; Mismatches 180; Indels 3; Gaps						
QY	241	CTGCCCCCGGCTTGGCGTGGGGGCTCTGGTGTGGGCTCTTTCGGGCTCTCTGGCAGCAT	300			
Db	95	CCGTCGCGCTCTGGCGCGCGCTGGCGCTCGACTGGAGCTCTGGCTGCGGCGCACGCT	154			
QY	301	CGAGCCCCAGGCGGTCCCTCCATATGCGTCGAGAACACAGACCTGCAGGGACCAAGAAA	360			
Db	155	TGCCCCCGCCAGGTGGGATTTACACCTAGCCCCCGAGCCCGGAGACATGCCGGCTCA	214			
QY	361	AGGAATACTATGAGCCCCAGCACCGCATCTGTGCTCTCCGCTGCGCGCGCACGCACTATG	420			


```

|||||
Db 215 GAGAACTATGACGACAGACAGCTCAGATGTGCTGCAGCAATGCTCGCGGCCAACATG 274
QY 421 TCTCAGCTAAATGATGACGACAGACAGATTTGTGCCACATGTGCCGAGAAATTCCT 480
Db 275 CAAAGTCTTCTGACCAAGACCTCGGACACCGTGTGACTCTGTGAGGACAGACAT 334
QY 481 ACAACGAGCACTGGAATACCTGACCATCTGCCAGCTGTGCCGCCCTGTGACCCAGTGA 540
Db 335 ACACCCAGCTCTGAACCTGTTCCCGAGTCTTGAGCTGTGGCTCCG---CTGTAGCT 391
QY 541 TGGCCCTCGAGGATGCCCCCTGACAAACGGAAGACCAACGATGCCGCTGCCAGC 600
Db 392 CTGACCCAGGTGGAACCTCAAGCCTGCACTCGGGAACAGAACCCGATCTGCACCTGCAGC 451
QY 601 CGGGAATGTTCTGTGC 616
Db 452 CCGCTGTGTACTGCGC 467

RESULT 14
AAQ10990
ID AAQ10990 standard; cDNA to mRNA; 1640 BP.
XX AC AAQ10990;
XX DT 24-MAY-1991 (first entry)
XX DE Partial sequence of hTNF-R clone 1.
XX KW Tumour necrosis factor receptor; immune response; inflammation;
XX KW cachexia; septic shock; ss.
XX OS Homo sapiens.
XX FH Key
XX FT CDS
XX FT mat_peptide
XX FT sig_peptide
XX PN EP418014-A.
XX PD 20-MAR-1991.
XX PF 10-SEP-1990; 90EP-0309875.
XX PR 10-MAY-1990; 90US-0523635.
XX PR 11-SEP-1989; 89US-0405370.
XX PR 13-OCT-1989; 89US-0421417.
XX FA (IMMU-) IMMUNEX CORP.
XX PI Smith CA, Goodwin RG, Beckmann PM;
XX WPI: 1991-082230/12.
XX P-PSDB: AAR11141.
XX New tumour necrosis factor -alpha and -beta receptors - and DNA
XX encoding these used to regulate immune responses in treatment of
XX cachexia, septic shock or side-effects of cytokine therapy.
XX Claim 1; Fig 2; 41pp; English.
XX The sequence was obtd. from a clone isolated from library prepd.
XX from a human fibroblast cell line, WI-26 VA4 (ATCC CCL 95.1).
XX The clone is deposited as Accession No. 68088 under the name
XX pCAV/NOT-TNF-R. The DNA can be truncated to produce sequences which
XX express soluble receptor comprising residues 1-235, 1-185 or 1-163
XX of the protein.
XX See also AAQ10991.

```

```

XX SQ Sequence 1640 BP; 328 A; 543 C; 502 G; 267 T; 0 other;
Query Match 3.6%; Score 72; DB 12; Length 1640;
Best Local Similarity 51.3%; Pred. No. 5.8e-07;
Matches 193; Conservative 0; Mismatches 180; Indels 3; Gaps 1;

QY 241 CTGCCCCCGGCTGCGCTGGGGGCTCTGGTGTGTGGGCTCTTTCGGGCTCTTGGCAGCAT 300
Db 95 CCGTCGCGCTGCGCGCGCTGGCCCTGCGACTGAGCTCTGGGCTGGGCGCACGCT 154
QY 301 CGCAGCCCCAGGCGGTGCCCTCATATGCTCGGAGAACACAGACCTCGAGGGACACGAAA 360
Db 155 TGCCTGCGCGAGGTGGCATTTACACCTTACGCTTACGCTGCGGAGCGCGGAGCACATGCGCGCTCA 214
QY 361 AGGAATACTATGAGCCCCAGCAGCATCTGCTCCGCTGCGCGCGGACCATATG 420
Db 215 GAGAACTATGACGACAGACAGCTCAGATGTGCTGCAGCAATGCTCGCGGCCAACATG 274
QY 421 TCTCAGCTAAATGATGACGACAGACAGATTTGTGCCACATGTGCCGAGAAATTCCT 480
Db 275 CAAAGTCTTCTGACCAAGACCTCGGACACCGTGTGACTCTGTGAGGACAGACAT 334
QY 481 ACAACGAGCACTGGAATACCTGACCATCTGCCAGCTGTGCCGCCCTGTGACCCAGTGA 540
Db 335 ACACCCAGCTCTGGAACCTGGGTTCGCGAGTCTTGTGCTGTGGTCCCG---CTGTAGCT 391
QY 541 TGGGCTCGAGGAGATTGCCCTGCGACAGCAACGGAAGACCAACGATGCCGCTGCCAGC 600
Db 392 CTGACCCAGGTGGAACCTCAAGCTTGCACATCGGGAACAGAACCCGATCTGCACCTGCAGC 451
QY 601 CGGGAATGTTCTGTGC 616
Db 452 CCGCTGTGTACTGCGC 467

RESULT 15
AAQ49931
ID AAQ49931 standard; cDNA to mRNA; 1641 BP.
XX AC AAQ49931;
XX DT 29-APR-1994 (first entry)
XX DE TNF-R cDNA.
XX KW Human; tumour necrosis factor receptor; TNF-R; interleukin-1 receptor;
XX KW IL-1R; fusion protein; linker; TNF; IL-1; cachexia; cerebral malaria;
XX KW rheumatoid arthritis; diabetes; multiple sclerosis; septic shock;
XX KW pulmonary fibrosis; silicosis; allograft; xenograft; rejection;
XX KW graft versus host disease; sepsis; inflammation; allergy;
XX KW autoimmune dysfunction; ss.
XX OS Homo sapiens.
XX FH Key
XX FT CDS
XX FT Location/Qualifiers
XX FT 88..1473
XX FT /*tag= a
XX FT /product= hTNF-R
XX FT sig_peptide
XX FT 88..153
XX FT /*tag= b
XX FT mat_peptide
XX FT 154..1470
XX FT /*tag= c
XX PN WO9319777-A.
XX PD 14-OCT-1993.
XX PF 26-MAR-1993; 93WO-US02938.
XX PR 30-MAR-1992; 92US-0860710.
XX FA (IMMU-) IMMUNEX CORP.

```



```
XX Smith CA;
PI
XX
XX WPI: 1993-336592/42.
DR P-PSDB; AAR42058.
XX
XX New fusion protein tumour necrosis factor and human interleukin-1
PT receptor - useful in therapy, diagnosis and assays of e.g.
PT rheumatoid arthritis, diabetes, cerebral malaria, sepsis, etc.
XX
XX Disclosure; Fig 2; 85pp; English.
XX
XX The sequences given in AAQ49931-32 encode human tumour necrosis factor
CC receptor (TNF-R) and the sequences in AAQ49933-34 encode human
CC interleukin-1 receptor (IL-1R). These sequences were used in the
CC production of a fusion protein which conformed to one of the
CC formulae:
CC TNF-R-linker-TNF-R-linker-IL-1R
CC IL-1R-linker-TNF-R-linker-TNF-R or
CC TNF-R-linker-TNF-R
CC The linker may comprise 5-100 amino acids selected from Gly, Asp,
CC Ser, Thr and Ala. These linkers separate the individual moieties
CC by such a distance that each component of the fusion protein is
CC capable of folding into the secondary or tertiary structure required
CC for its biological activity. These fusion proteins may be used in
CC therapy, diagnosis and assays for conditions mediated by TNF or IL-1,
CC particularly in conditions in which both TNF and IL-1 play a causative
CC role. They may be used to treat cachexia, rheumatoid arthritis,
CC diabetes, multiple sclerosis, pulmonary fibrosis and silicosis,
CC cerebral malaria, allograft and xenograft rejection in graft verses
CC host disease, sepsis, septic shock, inflammation, allergies and
CC autoimmune dysfunctions.
XX
XX Sequence 1641 BP; 328 A; 544 C; 502 G; 267 T; 0 other;
SQ
Query Match 3.6%; Score 72; DB 14; Length 1641;
Best Local Similarity 51.3%; Pred. No. 5.8e-07;
Matches 193; Conservative 0; Mismatches 180; Indels 3; Gaps 1;
QY 241 CTGCCCCCGGCTGGGCGCTCTGGTGTGGGCGCTTCGGGCTCCTGGCAGCAT 300
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 95 CCGTGGCGGTCTGGGCGGCGCTGGCGCTGGAGCTCTGGGCTGGCGGCGACGCT 154
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 301 CGCAGCCCCAGGCGGTGCTTCCATATGCGTGGGAGAACACAGACCTGCAGGGACCGAGAAA 360
Db || | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 155 TGCCCCCGCCAGGTGGCATTTACACCTACGCCCGGAGCCGGAGACATGCGGCTCA 214
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 361 AGGAATACTATGAGCCCGACGACCGCATGTGCTGCCCTGCCCGCCAGGACACCTATG 420
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 215 GAGAATACTATGACACAGACAGCTCAGATGTGCTGCAGAAATGCTCGCGGGCCAAACATG 274
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 421 TCTCAGCTAATGTAGCCGCATCGGGACACAGTTGTGCCACATGTGCCGAGAAATTCCT 480
Db || | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 275 CAAAAGTCTTCTGTACCAAGACCTCGGACACCGCTGTGTGACTCTGTGAGACAGCACAT 334
Db || | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 481 ACAACGAGCACTGAACTACCTGACCATGTGCCAGCTGTGCCCGCCCTGTGACCCAGTGA 540
Db || | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 335 ACACCCAGCTCTGGAAGTGGGTCCCGAGTGCTTGAGCTGTGGCTCCCG---CTGTAGCT 391
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 541 TGGGCTTCGAGAGATTGCCCTTCGACAAAGCAACGAAAGACCCAGTGCCTGCCAGC 600
Db || | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 392 CTGACCAAGGTGGAACACTCAAGCCTGCACCTCGGGAACAGAACCCGACATCGCACCTGCAGGC 451
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 601 CGGGAATGTTCTGTGC 616
Db || | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 452 CCGGCTGGTACTGCGC 467
```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:09:45 ; Search time 2289.32 Seconds
(without alignments)
14021.414 Million cell updates/sec

Title: US-09-917-372-2
Perfect score: 1982
Sequence: 1 gccccggcgccagctcgtc.....ggtatggggaagaggtttgg 1982

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpi:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: gb_gss:*
- 18: em_gss_hum:*
- 19: em_gss_inv:*
- 20: em_gss_pln:*
- 21: em_gss_vrt:*
- 22: em_gss_fun:*
- 23: em_gss_mam:*
- 24: em_gss_mus:*
- 25: em_gss_other:*
- 26: em_gss_pro:*
- 27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	861.2	43.5	954	14	BQ930450
2	821.6	41.5	882	14	BQ691803
3	818	41.3	933	9	AL564542
4	800	40.4	1006	14	BM923844
5	798.2	40.3	1109	13	BM463063
6	797.4	40.2	978	9	AL564472

7	794.8	40.1	993	9	AL513838	
8	794.2	40.1	816	13	BI908262	
9	785.6	39.6	867	13	BI768812	
10	755.4	38.1	840	13	BI906190	
11	746.6	37.7	954	13	BI821009	
12	729	36.8	895	14	BQ957322	
13	723.2	36.5	891	13	BI911211	
14	716.8	36.2	788	13	BI818683	
15	714.4	36.0	881	13	BI761249	
16	713.4	36.0	808	13	BI818411	
17	710.4	35.8	916	13	BI906596	
18	708.4	35.7	968	14	BQ653019	
19	704	35.5	917	13	BI760125	
20	698.6	35.2	932	13	BI765370	
21	696.4	35.1	801	12	BF792867	
22	695.2	35.1	1102	14	BM925105	
23	690.8	34.9	710	13	BI907277	
24	681.6	34.4	933	14	BQ214865	
c	25	679.8	34.3	801	9	AI193089
26	679.6	34.3	918	13	BI770354	
27	676.6	34.1	764	12	BG548459	
28	670.2	33.8	749	12	BF789430	
29	668.4	33.7	735	12	BG331666	
30	665.8	33.6	942	13	BI818552	
31	665	33.6	1239	14	BM923679	
32	664.4	33.5	666	14	BM767318	
33	660	33.3	732	9	AUI38624	
c	34	649.6	32.8	794	9	AI193092
35	638.4	32.2	780	9	AUI36112	
36	637.8	32.2	814	9	AUI18203	
37	636	32.1	636	14	BM849181	
c	38	633	31.9	742	14	BQ000587
39	632.6	31.9	968	14	BQ898015	
40	632	31.9	681	9	AUI39374	
41	630.4	31.8	920	12	BG026287	
42	629.2	31.7	634	14	BM746911	
43	626	31.6	841	12	BG750836	
44	622	31.4	622	14	BQ636409	
c	45	620.4	31.3	721	14	BM985373

ALIGNMENTS

RESULT 1
BQ930450
LOCUS BQ930450 954 bp mRNA linear EST 20-AUG-2002
DEFINITION AGENCOURT_8946139 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6462494
ACCESSION BQ930450
VERSION BQ930450.1 GI:22345481
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 954)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2652 row: h column: 15
High quality sequence stop: 576.
Location/Qualifiers
1. .954


```
Db 121 GGCACATCTACATCTACATGGACAGTACTGGGGGACACACCGGTCTCTGGAGACCTC 180
Qy 1245 CCAGCTACCCCCGAACTCCATACCCATCCCGAAGAGGGGAGCCCTGGCCCTCCCGGG 1304
Db 181 CCAGCTACCCCCGAACTCCATACCCATCCCGAAGAGGGGAGCCCTGGCCCTCCCGGG 240
Qy 1305 CTCTCTACACCCACCCAGGAGATGGCAAGCTTGGACCTAGCGGAGACAGACACTGT 1364
Db 241 CTCTCTACACCCACCCAGGAGATGGCAAGCTTGGCAAGCTTGGCAAGCTTGGCAAGCTGT 300
Qy 1365 GGTGCCACACCTCTAACAGGGGCCCAAGGAACCAATTATACCCATGACTGACTGAGT 1424
Db 301 GGTGCCACACCTCTAACAGGGGCCCAAGGAACCAATTATACCCATGACTGACTGAGT 360
Qy 1425 CTGAGAAAGCGAGAAGAGGGGGGCAAGAGGCACCTTCTCCCTTGAGGCTGCCCTGCC 1484
Db 361 CTGAGAAAGCGAGAAGAGGGGGGCAAGAGGCACCTTCTCCCTTGAGGCTGCCCTGCC 420
Qy 1485 CACGTGGGATTCACAGGGGCTGAGTAGAGGCCCGGGGAGGAGAGCCCTAAGGGATTAA 1544
Db 421 CACGTGGGATTCACAGGGGCTGAGTAGAGGCCCGGGGAGGAGAGCCCTAAGGGATTAA 480
Qy 1545 GCTCAGACACCTCTGAGAGCAGGTGGGCACCTGGGTAGGTGCGCTCCACAGGACTC 1604
Db 481 GCTCAGACACCTCTGAGAGCAGGTGGGCACCTGGGTAGGTGCGCTCCACAGGACTC 540
Qy 1605 TCCCTACTGCTGAGCAAACTGAGGCTCCCGCAGACACCCACCCCTGGGGTGC 1664
Db 541 TCCCTACTGCTGAGCAAACTGAGGCTCCCGCAGACACCCACCCCTGGGGTGC 600
Qy 1665 TCAGCCTCAGGCAGCGGACATGATACCAACTGCTGCCACTACGSCACGCCGCA 1724
Db 601 TCAGCCTCAGGCAGCGGACATGATACCAACTGCTGCCACTACGSCACGCCGCA 660
Qy 1725 CCGAGCAGCGCACCGAGGAGCGCCACACGGTCACCTGCAAGGAGTCAAGGGCCCT 1784
Db 661 CCGAGCAGCGCACCGAGGAGCGCCACACGGTCACCTGCAAGGAGTCAAGGGCCCT 720
Qy 1785 CTAAAGGATTGCT-GGTGCTCATCCCCAAGCTTCAGAGACCTTTGGGGTTCCACACTC 1843
Db 721 CTAAAGGATTGCTGGGTGCTCATCCCCAAGCTTCAGAGACCTTTGGGGTTCCACACTC 780
Qy 1844 ACGTGGAGTGGGTAGACCTGATGATGAATGAATATATA-GGAGAGCGCTCCTT--C 1899
Db 781 ACGTGGAGTGGGTAGACCTGATGATGAATGAATATATAAGGAGGAGCGCTCCTTCCC 840
Qy 1900 CCTCCCTCTCTAGAGGAGGAGGAGGAGTCAATTAACAA 1938
Db 841 TCCCTCTCTANNAGGAGAGAAAGGAGTATTANCA 879

RESULT 3
AL564542 933 bp mRNA linear EST 16-FEB-2001
LOCUS AL564542 LTI_NFL001_NBC4 Homo sapiens cdna clone CS0DM007YN10 5
DEFINITION prime, mRNA sequence.
ACCESSION AL564542
VERSION AL564542.1 GI:12915052
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 933)
AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
location/Qualifiers
1..933 source
```

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DM007YN10"
/clone_lib="LTI_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 163 a 335 c 257 g 166 t 12 others
ORIGIN

Query Match 41.3%; Score 818; DB 9; Length 933;
Best Local Similarity 97.9%; Pred. No. 6.9e-175;
Matches 826; Conservative 11; Mismatches 6; Indels 1; Gaps 1;

Qy 57 GAGCCCGGGCTGCGCGCTCCCGCCCTGGGGTGCACATCGGCCCTGAGTCCCTGCCAG 116
Db 1 GAGCCCGGGCTGCGCGCTCCCGCCCTGGGGTGCACATCGGCCCTGAGTCCCTGCCAG 60
Qy 117 GCTCTGGGCTCGGCAGCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCTTCT 176
Db 61 GCTCTGGGCTCGGCAGCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCTTCTCT 120
Qy 177 CCCAGGCCCGCCACCTTGTGGCCCGCTGGCGAGTGGCGGCATGCTCTCTCTTGGGCC 236
Db 121 CCCAGGCCCGCCACCTTGTGGCCCGCTGGCGAGTGGCGGCATGCTCTCTCTTGGGCC 180
Qy 237 ACCTCTCCCCCGCCCTGGCCCTGGGGCCCTGTGGTGTGGGCCCTTTCGGGCTCTTGSCA 296
Db 181 ACCTCTCCCCCGCCCTGGCCCTGGGGCCCTGTGGTGTGGGCCCTTTCGGGCTCTTGSCA 240
Qy 297 GCATCGCAGCCCGCAGCGGTGCTTCATATGCTGCGAGAACACAGACTCGCAGGAGCAG 356
Db 241 GCATCGCAGCCCGCAGCGGTGCTTCATATGCTGCGAGAACACAGACTCGCAGGAGCAG 300
Qy 357 GAAAGGAATACTATGAGCCCGCAGCAGCATGCTCTCCCGCTGCCCGCCAGGACCC 416
Db 301 GAAAGGAATACTATGAGCCCGCAGCAGCATGCTCTCCCGCTGCCCGCCAGGACCC 360
Qy 417 TATGCTCAGTAAATGAGCCGATCCGGGACACAGTTTGTGCCACATGTGCCGAGAA 476
Db 361 TATGCTCAGTAAATGAGCCGATCCGGGACACAGTTTGTGCCACATGTGCCGAGAA 420
Qy 477 TCCTACACGAGCACTGGAACCTACTGACCATCTGCCAGCTGTGCCGCCCTGTGACCCA 536
Db 421 WCMTAAMAGCACTGGAACCTACTGACCATCTGCCAGCTGTGCCGCCCTGTGACCCA 480
Qy 537 GTGATGGCCCTCGAGGAGATTGCCCTCGCAAGCAAAACGGAAGACCCAGTCCGCTGC 596
Db 481 GTGATGGCCCTCGAGGAGATTGCCCTCGCAAGCAAAACGGAAGACCCAGTCCGCTGC 540
Qy 597 CAGCGGGAATGTTCTGTCTGTCTGGGCCCTCGAGTGTACACACTGCGAGCTACTTCT 656
Db 541 CAGCGGGAATGTTCTGTCTGTCTGGGCCCTCGAGTGTACACACTGCGAGCTACTTCT 600
Qy 657 GACTGCCCGGCTGCGCACTGAAGCGAGCTCAAGATGAAGTTGGGAGGGTAAACACAC 716
Db 601 GACTGCCCGGCTGCGCACTGAAGCGAGCTCAAGATGAAGTTGGGAGGGTAAACACAC 660
Qy 717 TGGCTCCCTCGCAGGAGGCACTTCCAGAAATACCTCTCCCGCCAGCGCCGCTGCCAG 776
Db 661 TGGCTCCCTCGCAGGAGGCACTTCCAGAAATACCTCTCCCGCCAGCGCCGCTGCCAG 720
Qy 777 CCCACACAGGTGTGAGAACCAAGGTCTGTGTGGAGCAGCTCCAGGCACTGCCCGCAGTCC 836
```

```

Db 721 CCCACACACAGGTGTGAGAACCAAGTCTGGTGGAGGACAGTCCAGGCACTGCCAGTCC 780
QY 837 GACACAACTGCAAAAATCCA-TTAGAGGCACCTGCCCCAGAGATGTTCAGGATCGTCTGCT 895
Db 781 GACACAACTGCAAAAATCCATTAGAGGCACCTGCCCCAGAGATGTTCAGGAGCAATGCT 840
QY 896 CAAG 899
Db 841 GATG 844

RESULT 4
BM923844
LOCUS
DEFINITION
  AGENCOURT_6709962 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5759970
  5', mRNA sequence.
ACCESSION
  BM923844
VERSION
  BM923844.1 GI:19374223
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 1006)
  NIH-MGC http://mgi.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: Life Technologies, Inc.
  cDNA Library Preparation: Life Technologies, Inc.
  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: Agencourt Bioscience Corporation
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov
  Plate: LLAM12806 row: d column: 19
  High quality sequence stop: 688.
FEATURES
  Location/Qualifiers
    1..1006
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="IMAGE:5759970"
      /clone_lib="NIH_MGC_116"
      /lab_host="DH10B"
      /note="Organ: pooled colon, kidney, stomach; Vector:
      pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
      source anonymous pool of 3 colons, age 26 yo male, 49 yo
      female, 71 yo male colon; 46 yo male kidney, and pool of 2
      stomachs, 62 yo male and 70 yo female. Library is
      oligo-dT primed and directionally cloned (EcoRV site is
      destroyed upon cloning). Average insert size 1.4 kb,
      insert size range 1-3 kb. Library is normalized and
      enriched for full-length clones and was constructed by C.
      Gruber (Invitrogen). Research Genetics tracking code
      023. Note: this is a NIH_MGC Library."
BASE COUNT
  186 a 364 c 276 g 180 t
ORIGIN
  Query Match
  Best Local Similarity 40.4%; Score 800; DB 14; Length 1006;
  Matches 823; Conservative 0; Mismatches 20; Indels 1; Gaps 1;
  QY 57 GAGGCGCGCGCTGGCGCTCCCGGCGCTGGGTGCACATCGCGCCCTGAGTCCCGTCCAG 116
  Db 1 GAGGCGCGCGCTGGCGCTCCCGGCGCTGGGTGCACATCGCGCCCTGAGTCCCGTCCAG 60
  QY 117 GCTCTGGGCTCGGAGCGCGCGCACCGCTGCCAGGACGTGCGGCGCTCTCTGCTCTCT 176
  Db 61 GCTCTGGGCTCGGAGCGCGCGCACCGCTGCCAGGACGTGCGGCGCTCTCTGCTCTCT 120
  QY 177 CCCAGGCCCCACAGTGTGCTGGCGCGCTGGCGAGTGGCGCGCATGCTCTGCTGGGCC 236

```

```

Db 121 CCAGGCCCCACAGTTGCTGGCGCGCTGGCGAGTGGCGCCATGCTCTGCTGCTGGCGC 180
QY 237 ACCTCTGCCCCCGGCTGCTGGCGCGCTGGCGAGTGGCGCTCTGCTGGCGCTCTGCTGGCA 296
Db 181 ACCTCTGCCCCCGGCTGCTGGCGCGCTGGCGAGTGGCGCTCTGCTGGCGCTCTGCTGGCA 240
QY 297 GCATCGACGCCCGAGCGGTGCTCCATATGCGTCCGAGAACACAGACCTGCAGGGACCCAG 356
Db 241 GCATCGACGCCCGAGCGGTGCTCCATATGCGTCCGAGAACACAGACCTGCAGGGACCCAG 300
QY 357 GAAAGGAATACTATGAGCCCCAGCACCGCATCTGCTGCTCCCGTGGCGCGCCAGGACCC 416
Db 301 GAAAGGAATACTATGAGCCCCAGCACCGCATCTGCTGCTCCCGTGGCGCGCCAGGACCC 360
QY 417 TATGCTCAGCTAAATGTAGCGCATCCGGGACACAGTTTGTGCCACATGTGCCGAGAT 476
Db 361 TATGCTCAGCTAAATGTAGCGCATCCGGGACACAGTTTGTGCCACATGTGCCGAGAT 420
QY 477 TCCTACACGAGCAGCTGGAACTACCTGACCATCTGCCAGCTGTGCCGCCCTGTGACCCA 536
Db 421 TCCTACACGAGCAGCTGGAACTACCTGACCATCTGCCAGCTGTGCCGCCCTGTGACCCA 480
QY 537 GTGATGGGCTCGAGGAGATGCCGCCCTGCACAAGCAAAACGGAAGACCCAGTGCCTGTC 596
Db 481 GTGATGGGCTCGAGGAGATGCCGCCCTGCACAAGCAAAACGGAAGACCCAGTGCCTGTC 540
QY 597 CAGCGGGGAATCTGCTGCTGCTGGCGCGCTGGCGAGTGTACACACTGCGAGCTACTTCT 656
Db 541 CAGCGGGGAATCTGCTGCTGCTGGCGCGCTGGCGAGTGTACACACTGCGAGCTACTTCT 600
QY 657 GACTGCCGCCCTGGCAGCTGAAGCCGAGCTCAAGATGAAGTTGGGAAGGTAAACACCCAC 716
Db 601 GACTGCCGCCCTGGCAGCTGAAGCCGAGCTCAAGATGAAGTTGGGAAGGTAAACACCCAC 660
QY 717 TCGTCTCCCTGCAAGCGAGGCACTTCCAGATACCTCTCCCGCAGCGCCGCTGCGCAG 776
Db 661 TCGTCTCCCTGCAAGCGAGGCACTTCCAGATACCTCTCCCGCAGCGCCGCTGCGCAG 720
QY 777 CCCCACACAGGTGTGAG-AACCAAGGTCTGGTGGAGGAGCTCCAGGACCTGCCAGCTC 835
Db 721 CCCCACACAGGTGTGAGAACCAAGTCTGGTGGAGGAGCTCCAGGACCTGCCAGCTC 780
QY 836 CGACACAACTGCAAAAATCCATTAGAGCCACTGCCCGCAGAGATGTCCAGGATCGTCTGCT 895
Db 781 CGACACAACTGCAAAAATCCATTAGAGCCACTGCCCGCAGAGATGTCCAGGACCTGCT 840
QY 896 CAAG 899
Db 841 GATG 844

RESULT 5
BM463063
LOCUS
DEFINITION
  AGENCOURT_6447515 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5537750
  5', mRNA sequence.
ACCESSION
  BM463063
VERSION
  BM463063.1 GI:18512103
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 1109)
  NIH-MGC http://mgi.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: ATCC/DCTD/DTF
  cDNA Library Preparation: Life Technologies, Inc.
  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

```

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MCC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: L1AM12229 row: i column: 15
High quality sequence stop: 659.
Location/Qualifiers

FEATURES
source

1. .1109
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5537750"
/clone_lib="NIH_MCC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."

BASE COUNT 211 a 394 c 309 g 195 t
ORIGIN

Query Match 40.3%; Score 798.2; DB 13; Length 1109;
Best Local Similarity 98.1%; Pred. No. 2.2e-170;
Matches 839; Conservative 0; Mismatches 13; Indels 3; Gaps 3;

QY 28 ACTTCTGAGCTCCGCGCATGGAGCCCTGGAGCGCGGCTGGCGCTCCCGGCCCTGGG 87
DB 2 ACTTCTGAGCTCCGCGCATGGAGCCCTGGAGCGCGGCTGGCGCTCCCGGCCCTGGG 61
QY 88 GTGCACATCGGCGCTGAGTCCGCTCCAGGCTCTGGGCTGGGCGAGCGCGCCACCGT 147
DB 62 GTGCACATCGGCGCTGAGTCCGCTCCAGGCTCTGGGCTGGGCGAGCGCGCCACCGT 121
QY 148 GCCAGGAGCTGGGCTCTGCTCCCTCCCTCCAGGCGCCAGCTGGCTGGCGCGCTGGCC 207
DB 122 GCCAGGAGCTGGGCTCTGCTCCCTCCCTCCAGGCGCCAGCTGGCTGGCGCGCTGGCC 181
QY 208 GAGTGGCGCCCATGCTCTGCTTGGGCGACCTCTGCCCCCGGCTGGGCTGGGCGCTC 267
DB 182 GAGTGGCGCCCATGCTCTGCTTGGGCGACCTCTGCCCCCGGCTGGGCTGGGCGCTC 241
QY 268 TGGTGGTGGGCTCTTGGGCTCTGGGCGAGATCGGAGCGCCAGCGGCTGCTCCATATG 327
DB 242 TGGTGGTGGGCTCTTGGGCTCTGGGCGAGATCGGAGCGCCAGCGGCTGCTCCATATG 301
QY 328 CGTGCGAGAACAGACCTCGAGGACCGAGAAAGAAATATAGTATGAGCCCGACCGCA 387
DB 302 CGTGCGAGAACAGACCTCGAGGACCGAGAAAGAAATATAGTATGAGCCCGACCGCA 361
QY 388 TGTGCTGCTCCGCTGCGCGCCAGGACCTATGCTCAGCTAAATGATAGCCGATCCGGG 447
DB 362 TGTGCTGCTCCGCTGCGCGCCAGGACCTATGCTCAGCTAAATGATAGCCGATCCGGG 421
QY 448 ACACAGTTTGTGCCACATGTCGGGATCTGGGCGAGATCTTCTACACGAGCACTGGA 507
DB 422 ACACAGTTTGTGCCACATGTCGGGATCTGGGCGAGATCTTCTACACGAGCACTGGA 481
QY 508 TCTGCGAGCTGTGCGCGCTGTGACCCAGTGTGAGGCTCGAGGAGATTGCCCGCTGCA 567
DB 482 TCTGCGAGCTGTGCGCGCTGTGACCCAGTGTGAGGCTCGAGGAGATTGCCCGCTGCA 541
QY 568 CAAGCAAAACGAGACCGAGTCCGCTGCCAGCGCGGAAATGTTCTGTGTGCTGGGCGC 627
DB 542 CAAGCAAAACGAGACCGAGTCCGCTGCCAGCGCGGAAATGTTCTGTGTGCTGGGCGC 601
QY 628 TCGAGTGTACACATCGGAGCTACTTCTGACTGCCCGCTGGCACTGAAGCGGAGCTCA 587
DB 602 TCGAGTGTACACATCGGAGCTACTTCTGACTGCCCGCTGGCACTGAAGCGGAGCTTC 661
QY 688 AAGATGAAGTGGGAGGGTAAACAACTGCTGCCCTGCAAGGCGAGGCGACTTCCAGA 747
DB 662 AAGATGAAGTGGGAGGGTAAACAACTGCTGCCCTGCAAGGCGGCGACTTCCAGA 721

QY 748 ATACCTCTCTCCCGAGCGCGCTGCCAGCCCGACACACAGGTGTGAG-AACCAAGGTCTG 806
DB 722 ATACCTCTCTCCCGAGCGCGCTGCCAGCCCGACACACAGGTGTGAGAAACAGGCGCTG 781
QY 807 GTGGAGGAGCTCCAGGCACT-GCCAGTCCGACACAACT-GCAAAATCCATTAGAC 864
DB 782 GTGGAGGAGCTCCAGGCACTGGCGAGTCCGACACAACTGGCAAAATCCCTTAGAG 841
QY 865 CACTGCCCCCAGAGA 879
DB 842 CCTGGCCCCCAGAGA 856
AL564472 978 bp mRNA linear EST 16-FEB-2001
AL564472 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DM007N10 3
prime, mRNA sequence.
AL564472
AL564472.1 GI:12914912
EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 978)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1. .978
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DM007N10"
/clone_lib="LTI_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com".
BASE COUNT 160 a 280 c 305 g 227 t 6 others
ORIGIN
Query Match 40.2%; Score 797.4; DB 9; Length 978;
Best Local Similarity 95.8%; Pred. No. 3.3e-170;
Matches 932; Conservative 5; Mismatches 21; Indels 15; Gaps 11;
QY 1011 GTTCTCCCATATCCACTGGGCTCCCGCAGCGCCAGTGTGGAGCAGGGGTGCCGCAA 1070
DB 974 GTTCTCCCATATCCACT-GGCTCCCGCCAGCGCCCAKTTT--GGAGCGAGGGTGCCGCAA 918
QY 1071 CAGCAGAGTCTCTGGACCTGACAGGAGCGCAGTTGGAACCGGGGAGCAGAGCCAG 1130
DB 917 CAGCAGAGTCTCT-GACCTGACCA-GGAGCGCAGCTKGAACCC--GGAGCAGAGCCAA- 863
QY 1131 GTGGCCACAGGTACCAATGGCAATTCATGTCACCGCGGGTCTATGACTATCACTGGCAAC 1190
DB 862 GTGGCCACAGGTACCAATGGCAATTCATGTCACCGCGGG--TCTATATCACT-GCAAC 806
QY 1191 ATCTACATCTACAAATGACCACTGCGGGGACACCGGGTCTCTGGAGACCTCCAGCT 1250
|||||

Db 805 ATCTACATCTACATGGACCAAGWACT - GGGGAGACACCGGGTCTCTGGAGACCTCCAGCT 747

Qy 1251 ACCCCCGAACCTCCATACCCCATCCCGAAGAGGGAGCCCTGGCCCTCCCGGGCTCTCT 1310

Db 746 ACCCCCGAACCTCCATACCCCATCCCGAAGAGGG - ACCTGSCCTCCCGGGCTCTCT 689

Qy 1311 ACACCCCAACAGGAGAGTGGCAAGCTTTGGCACTTAGCGGAGACAGACACTGTGGTGC 1370

Db 688 ACACCCCAACAGGAGAGTGGCAAGCTTTGGCACTTAGCGGAGACAGACACTGTGGTGC 629

Qy 1371 ACACCCCTACAGGGGCGCCCAAGAACCAATTTATCACCCATGACTGACTGAGTCTGAG - 1429

Db 628 ACACCCCTACAGGGGCGCCCAAGAACCAATTTATCACCCATGACTGAGGAGTCTGAGT 569

Qy 1430 AAAAGGCAGAAAGAGGGGGGCACAAAGGCACCTTCTCCCTTGAGGCTGCCCTGCCACGT 1489

Db 568 AAAAGGCAGAAAGAGGGGGGCACAAAGGCACCTTCTCCCTTGAGGCTGCCCTGCCACGT 509

Qy 1490 GGGATTACAGGGGCTGAGTAGGGCCCGGGGAAGACAGAGCCCTAAAGGGATTAAAGGCTCA 1549

Db 508 GGGATTACAGGGGCTGAGTAGGGCCCGGGGAAGACAGAGCCCTAAAGGGATTAAAGGCTCA 449

Qy 1550 GACACCTCTGAGACAGTGGSCACTGGCTGAGTACGCTGCCCTCCACAGGACTTCCTCT 1609

Db 448 GACACCTCTGAGACAGTGGSCACTGGCTGAGTACGCTGCCCTCCACAGGACTTCCTCT 389

Qy 1610 ACTGCTGAGCAAACTGAGGCTCCCGGCAGACCCACCCCTCCCTGGGCTGCTCAGC 1669

Db 388 ACTGCTGAGCAAACTGAGGCTCCCGGCAGACCCACCCCTCCCTGGGCTGCTCAGC 329

Qy 1670 CTCAGGCAGACAGGGGACATGATACCACTGTGCTGCCACTACAGACCGCCGACCGGA 1729

Db 328 CTCAGGCAGACAGGGGACATGATACCACTGTGCTGCCACTACAGGCACCGCCGACCGGA 269

Qy 1730 GCAGGCAACCGAGGAGCGCCACACAGCTACCTGCAAGGAGCTACAGGCGCCCTCTAAA 1789

Db 268 GCAGGCAACCGAGGAGCGCCACACAGCTACCTGCAAGGAGCTACAGGCGCCCTCTAAA 209

Qy 1790 GGATTCGTGGTCTCATCCCCCAAGCTTCAGAGACCTTTGGGGTTCCACACTTACGCTGG 1849

Db 208 GGATTCGTGGTCTCATCCCCCAAGCTTCAGAGACCTTTGGGGTTCCACACTTACGCTGG 149

Qy 1850 ACTGAGTAGACCTGCATGATGAAGTAAATATAGGAGGAGCGCTCTTCCCTCCCTCC 1909

Db 148 ACTGAGTAGACCTGCATGATGAAGTAAATATAGGAGGAGCGCTCTTCCCTCCCTCC 89

Qy 1910 TAGAGGAGGAAAGGAGTCAATTAACAACCTAGGGGTTGGTAGGATTCTTAGTATGG 1969

Db 88 TAGAGGAGGAAAGGAGTCAATTAACAACCTAGGGGTTGGTAGGATTCTTAGTATGG 29

Qy 1970 GGAAGAGTTTGG 1982

Db 28 GGAAGAGTTTGG 16

RESULT 7

AL513838

LOCUS

DEFINITION AL513838 LTI_NFL006_P12 Homo sapiens cDNA clone CL0BA002ZF12 5

prime, mRNA sequence.

ACCESSION AL513838

VERSION AL513838.1 GI:12777332

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 993)

AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1. 993

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CL0BA002ZF12"

/clone_lib="LTI_NFL006_P12"

/tissue_type="placenta"

/note="Vector: pCMVSPORT 6; Site: 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax: (1) 301 610 8371 Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com"

BASE COUNT 184 a 348 c 278 g 180 t 3 others

ORIGIN

Query Match 40.1%; Score 794.8; DB 9; Length 993;

Best Local Similarity 97.9%; Pred. No. 1.3e-169;

Matches 805; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 78 CGGCGCTGGGGTGACATCGGCGCTGAGTCCCGTCCAGGCTCTGGGCTCGGGCAGCGC 137

Db 5 CGGTCCGGAATTCGCGGATGCTGCTGAGTCCCGTCCAGGCTCTGGGCTCGGGCAGCGC 64

Qy 138 CGCACCGCTCGCCAGGACGTGCGGGCTTCCTGCTTCCTCCAGGCCCCACAGTTGCTGG 197

Db 65 CGCACCGCTCGCCAGGACGTGCGGGCTTCCTGCTTCCTCCAGGCCCCACAGTTGCTGG 124

Qy 198 CCGCTGCGCCAGTGGCGCATGCTGCTGCTTGGGCCACTCTGCCCCCGGCGTGGC 257

Db 125 CCGCTGCGCCAGTGGCGCATGCTGCTGCTTGGGCCACTCTGCCCCCGGCGTGGC 184

Qy 258 TGGGGGCTCTGGTCTGGGCTCTTCCGGCTCTGCGCAGCATGCGAGCCCGCAGCGGTG 317

Db 185 TGGGGGCTCTGGTCTGGGCTCTTCCGGCTCTGCGCAGCATGCGAGCCCGCAGCGGTG 244

Qy 318 CCTCATATGCTCGGAGAACACAGCTGCGAGGACAGGAAAGAAATATATGAGGCC 377

Db 245 CCTCATATGCTCGGAGAACACAGCTGCGAGGACAGGAAAGAAATATATGAGGCC 304

Qy 378 CAGCACGCATCTGCTGCTCCGCTGCCCGCAGGACCTATGCTCAGCTAAATGTAGC 437

Db 305 CAGCACGCATCTGCTGCTCCGCTGCCCGCAGGACCTATGCTCAGCTAAATGTAGC 364

Qy 438 CGCATCCGGGACACAGTTTGTGCCACATGTGCCGAGAATTCCTACAACGAGCAGTGAAC 497

Db 365 CGCATCCGGGACACAGTTTGTGCCACATGTGCCGAGAATTCCTACAACGAGCAGTGAAC 424

Qy 498 TACCTGACCATCTGCCAGCTGTCCGCGCTGTGACCCAGTGTATGGGCTTCGAGGAGAT 557

Db 425 TACCTGACCATCTGCCAGCTGTCCGCGCTGTGACCCAGTGTATGGGCTTCGAGGAGAT 484

Qy 558 GCCCCTGACACAGCAACGAGGAGGAGTCCGCTGCCAGCGGGATGCTTCTGTGCT 617

Db 485 GCCCCTGACACAGCAACGAGGAGGAGTCCGCTGCCAGCGGGATGCTTCTGTGCT 544

Qy 618 GCCTGGGCGCTCGAGTGTACACACTGCGAGCTACTTTCTGACTGCCCGCTGGCACTGAA 677

Db 545 GCCTGGGCGCTCGAGTGTACACACTGCGAGCTACTTTCTGACTGCCCGCTGGCACTGAA 604

Qy 678 GCCGAGCTCAAAGATGAAGTTGGGAAGGTAACAACTCGTCCCTCCGAAAGCAGGG 737

Db 605 GCCGAGCTCAAAGATGAAGTTGGGAAGGTAACAACTCGTCCCTCCGAAAGCAGGG 664

Qy 738 CACTTCCAGTAATCTCTCTCCCGCAGCGCGCTGCCAGCCCCACACAGGTGTGAGAAC 797

Db 665 CACTTCCAGTAATCTCTCTCCCGCAGCGCGCTGCCAGCCCCACACAGGTGTGAGAAC 724


```
QY 798 CAAGGTCTGGTGGAGGACGCTCAGGACACTGCCAGTCCGACACACCAACCTGCAAAATCCA 857
Db 725 CAAGGTCTGGTGGAGGACGCTCCAGGACTGCCAGTCCGACACCAACCTGCAAAATCCA 784
QY 858 TTAGAGCCACTGCCCCAGAGATGTGAGGATGCCTGCTCAAG 899
Db 785 TTAGAGCCACTGCCCCAGAGATGTGAGGATGCCTGCTGATG 826

RESULT 8
BI908262 816 bp mRNA linear EST 16-OCT-2001
LOCUS 603066987F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5216085 5',
DEFINITION mRNA sequence.
ACCESSION BI908262
VERSION BI908262.1 GI:16171181
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 816)
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL).
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11542 row: n column: 22
High quality sequence stop: 813.
Location/Qualifiers
1..816
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5216085"
/clone_lib="NIH_MGC_118"
/tissue_type="leukocyte"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV
(destroyed); RNA source leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dr primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 027. Note:
this is a NIH_MGC Library."
BASE COUNT 150 a 299 c 231 g 136 t
ORIGIN
Query Match 40.1%; Score 794.2; DB 13; Length 816;
Best Local Similarity 99.5%; Pred. No. 1.6e-169;
Matches 807; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 63 CGGCTGGCCGCTCCGGCCCTGGGTGCACATCGGCCCTGATCGCCGTCCTCCAGGCTCTG 122
Db 1 CGGCTGGCCGCTCCGGCCCTGGGTGCACATCGGCCCTGATCGCCGTCCTCCAGGCTCTG 60
QY 123 GGTCTGGGAGCGCGCCGACCGCTGCCAGGACGTGGGCTCTCTGCTTCTCTCCAGG 182
Db 61 GGTCTGGGAGCGCGCCGACCGCTGCCAGGACGTGGGCTCTCTGCTTCTCTCCAGG 120
QY 183 CCCCCACGTTGCTGGCGCTGGCGGAGTGCGGCATCTCTGCTGGGCGCACCTCT 242
Db 121 CCCCCACGTTGCTGGCGCTGGCGGAGTGCGGCATCTCTGCTTGGGCGCACCTCT 180

FEATURES
source
```

```
QY 243 GCCCCCGGCTGGCGCTCTGTGCTGGGCTCTTCTGGGCTCTTGGCAGCATCG 302
Db 181 GCCCCCGGCTGGCGCTCTGTGCTGGGCTCTTCTGGGCTCTTGGCAGCATCG 240
QY 303 CAGCCCCAGGCGGTGCTCCATATGCGTCGGAGAACCAACCTGAGGACCCAGGAAAG 362
Db 241 CAGCCCCAGGCGGTGCTCCATATGCGTCGGAGAACCAACCTGAGGACCCAGGAAAG 300
QY 363 GAATACTATGAGCCCCAGCACCGCATCTGCTGCTCCGCTGCCGCGCAGCACCTATGTC 422
Db 301 GAATACTATGAGCCCCAGCACCGCATCTGCTGCTCCGCTGCCGCGCAGCACCTATGTC 360
QY 423 TCAGCTAAATGTAGCCGATCCGGGACACAGTTTGTGCCACATGTGCCGAGAAATTCCTAC 482
Db 361 TCAGCTAAATGTAGCCGATCCGGGACACAGTTTGTGCCACATGTGCCGAGAAATTCCTAC 420
QY 483 AACGAGCACTGGAACCTGACCATCTGCCAGCTGTGCCGCCCTGTGACCCAGTGATG 542
Db 421 AACGAGCACTGGAACCTGACCATCTGCCAGCTGTGCCGCCCTGTGACCCAGTGATG 480
QY 543 GGCCTCGAGGAGATGGCCCTTGCACAAAGCAAGCAAGCAAGTGGCGCTGCCAGCCG 602
Db 481 GGCCTCGAGGAGATGGCCCTTGCACAAAGCAAGCAAGCAAGTGGCGCTGCCAGCCG 540
QY 603 GGAATGTTCTGTGCTGCTGGCCCTCGAGTGTACACACTGCGAGCTACTTTCTGACTGC 662
Db 541 GGAATGTTCTGTGCTGCTGGCCCTCGAGTGTACACACTGCGAGCTACTTTCTGACTGC 600
QY 663 CGCCTGGCACTGAAGCCGAGCTCAAGATGAAGTTGGGAAGGTAACCAACCTGCGTC 722
Db 601 CGCCTGGCACTGAAGCCGAGCTCAAGATGAAGTTGGGAAGGTAACCAACCTGCGTC 660
QY 723 CCTCTCAAGGAGGCACTTCCAGAACTCTCTCCGCCAGCGCCGCTGCCAGCCCCAC 782
Db 661 CCTCTCAAGGAGGCACTTCCAGAACTCTCTCTCCGCCAGCGCCGCTG-CAGCCCCAC 719
QY 783 ACCAGTGTGAGAACCAAGGTCTGTGGAGGAGCTTCCAGGCACTGCCAGTCCGACACA 842
Db 720 ACCAGTGTGAGAACCAAGGTCTGTGGAGGAGCTTCCAGGCACTGCCAGTCCGACACA 779
QY 843 ACCTCAAAATCCATTAGAGCCACTGCCCC 873
Db 780 AACTCAAAATCCATTAGAGCCACTGCCCC 810

RESULT 9
BI768812 867 bp mRNA linear EST 25-SEP-2001
LOCUS 603057713F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5207139 5',
DEFINITION mRNA sequence.
ACCESSION BI768812
VERSION BI768812.1 GI:15760390
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 867)
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11519 row: j column: 04
High quality sequence stop: 818.
Location/Qualifiers
```

FEATURES

```

source
1. .867
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5207139"
/clone_lib="NIH_MGC_122"
/lab_host="DH10B"
/notes="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC Library."
BASE COUNT 171 a 306 c 248 g 141 t 1 others
ORIGIN

Query Match 39.6%; Score 785.6; DB 13; Length 867;
Best Local Similarity 98.1%; Pred. No. 1.5e-167;
Matches 794; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 91 CACATCGCCCTGAGTCCCGTCCAGGCTCTGGGCTCGGCAGCCGCCACCGCTGCC 150
DB 1 CACATCGCCCTGAGTCCCGTCCAGGCTCTGGGCTCGGCAGCCGCCACCGCTGCC 60
QY 151 CAGGACGTGGGCTCTGCTTCTCCAGGCCCCACGTTGTCGCCGCTGGCCGAG 210
DB 61 CAGGACGTGGGCTCTGCTTCTCCAGGCCCCACGTTGTCGCCGCTGGCCGAG 120
QY 211 TGGCCGCACTCTCTGCTTGGGCCACCTCTGCCCCGGCTGGCTGGGGCTCTGG 270
DB 121 TGGCCGCACTCTCTGCTTGGGCCACCTCTGCCCCGGCTGGCTGGGGCTCTGG 180
QY 271 TGCTGGGCTCTTGGGCTCTTGCGACATCGACGCCCCAGGGGCTGCTCCATATCGT 330
DB 181 TGCTGGGCTCTTGGGCTCTTGCGACATCGACGCCCCAGGGGCTGCTCCATATCGT 240
QY 331 CGGAGACCACTGTCAGGACCCAGGAAAGAAATGTAATGACCCCGCCAGCCGATCT 390
DB 241 CGGAGACCACTGTCAGGACCCAGGAAAGAAATGTAATGACCCCGCCAGCCGATCT 300
QY 391 GCTGCTCCGCTGCGCCAGGACCTATGCTCAGCTAAATGATAGCGGATCGGGACA 450
DB 301 GCTGCTCCGCTGCGCCAGGACCTATGCTCAGCTAAATGATAGCGGATCGGGACA 360
QY 451 CAGTTTGTGCCATGTGCCGAGAAATTCCTACACGAGCACTGGAATACCTGACCATCT 510
DB 361 CAGTTTGTGCCATGTGCCGAGAAATTCCTACACGAGCACTGGAATACCTGACCATCT 420
QY 511 GCCAGCTGCGCCCTGTGACCCAGTGTAGGGCTCGAGGAGATGCCCCCTGCACAA 570
DB 421 GCCAGCTGCGCCCTGTGACCCAGTGTAGGGCTCGAGGAGATGCCCCCTGCACAA 480
QY 571 GCAACGGAAGACCACTGCGCTGCCAGCGGAAATGTTCTGCTGCTGGGCCCTCG 630
DB 481 GCAACGGAAGACCACTGCGCTGCCAGCGGAAATGTTCTGCTGCTGGGCCCTCG 540
QY 631 AGTGTACACATGGAGCTACTTTCTGACTGCCCGCTGGGACCTGAAGCCGAGCTCAAG 690
DB 541 AGTGTACACATGGAGCTACTTTCTGACTGCCCGCTGGGACCTGAAGCCGAGCTCAAG 600
QY 691 ATGAAGTTGGGAAGGTAACAAACACTGCGTCCCTGCAAGGCAAGGCACTTCCAGAATA 750
DB 601 ATGAAGTTGGGAAGGTAACAAACACTGCGTCCCTGCAAGGCAAGGCACTTCCAGAATA 660
QY 751 CCTCTCCCCAGGCCCGCTGCGAGCCCCACACAGGTGTGAGAACCAAGGTCTGTGG 810
DB 661 CCTCTCCCCAGGCCCGCTGCGAGCCCCACACAGGTGTGAGAACCAAGGTCTGTGG 720
QY 811 AGGAGCTCCAGGCACTGCCAGTCCGACACAACTGCAAAATCCATTTAGAGCACTGC 870
DB 911 AGGAGCTCCAGGCACTGCCAGTCCGACACAACTGCAAAATCCATTTAGAGCACTGC 870

```

Db 721 AGCAGCTCCAGCACTGGCCAGTCCGACACGAACTGCAAAATCCATTAGACCAACTG 780
 QY 871 CCCAGAGATGTCAGGATGCTGCTCAAG 899
 Db 781 CCCAGAGATGTCAGAACTGGCCTCAAG 809

RESULT 10
 BI906190 840 bp mRNA linear EST 16-OCT-2001
 LOCUS 603062886F1.NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5211812 5',
 DEFINITION mRNA sequence.
 ACCESSION BI906190
 VERSION BI906190.1 GI:16168845
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 840)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLML1531 row: 1 column: 21
 High quality sequence stop: 776.
 Location/Qualifiers
 source
 1. .840
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5211812"
 /clone_lib="NIH_MGC_118"
 /tissue_type="leukocyte"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV
 (destroyed); RNA source leukocytes from anonymous pool of
 non-activated adult donors. Library is oligo-dT primed
 and directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb, insert size range
 1.2-3.3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 027. Note:
 this is a NIH_MGC Library."
 BASE COUNT 155 a 300 c 246 g 138 t 1 others

Query Match 38.1%; Score 755.4; DB 13; Length 840;
 Best Local Similarity 97.4%; Pred. No. 1e-160;
 Matches 821; Conservative 0; Mismatches 16; Indels 6; Gaps 5;

QY 47 GGGAGCCCTGGAGCCCGGCTGGCGCTCCCGCCCTGGGGTGCACATCGGCCCTGAGT 106
 DB 1 GGGAGCCCTGGAGCCCGGCTGG-CGCTCCCGCCCTGGGGTGCACATCGGCCCTGAGT 59
 QY 107 CCGTCCCAAGCTCTGGGCTCGGCAGCCGCCACCGTCCGCCAGGAGTCTGGGGCTC 166
 DB 60 CCGTCCCAAGCTCTGGGCTCGGCAGCCGCCACCGTCCGCCAGGAGTCTGGGGCTC 119
 QY 167 CTGCTCTTCCCAAGCCCGCCACGTTGTCGCCCTGGCGGAGTGGCCCATGCTCT 226
 DB 120 CTGCTCTTCCCAAGCCCGCCACGTTGTCGCCCTGGCGGAGTGGCCCATGCTCT 179
 QY 227 GCCTTGGGCACTCTGCCCCCGGCTGGCTGGGGGCTCTGTGTGTGGGCTCTTCGG 286
 DB 180 GCCTTGGGCACTCTGCCCCCGGCTGGCTGGGGGCTCTGTGTGTGGGCTCTTCGG 239

FEATURES		Location/Qualifiers	
Source		1. .954	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:5176935"	
		/clone_lib="NIH_MGC_115"	
		/lab_host="DH10B"	
		/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dr primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research genetics tracking code 021. Note: this is a NIH_MGC Library."	
BASE COUNT		212 a 313 c 281 g 148 t	
ORIGIN			
Query Match		37.7%; Score 746.6; DB 13; Length 954;	
Best Local Similarity		93.3%; Pred. No. 1e-158;	
Matches		861; Conservative 0; Mismatches 44; Indels 18; Gaps 7;	
QY	885	GGATCGCTGCTCAAGAGGGCTCCGAGGAGGAGGAGCCCAATCCTGTAGCTGGAAGCTGG	944
Db	14	GGATCGCTGCTCAAGAGGGCTCCGAGGAGGAGGAGCCCAATCCTGTAGCTGGAAGCTGG	73
QY	945	GAGCTCCGAAGGCCATCCATCTCCCTGACTTGTGTAGCCACTGTACCCATTCT	1004
Db	74	GAGCTCCGAAGGCCATCCATCTCCCTGACTTGTGTAGCCACTGTACCCATTCT	133
QY	1005	GGAGATGTTTCCCGAGTATCCACTGGCTCCCGCAGCCCGCCAGTTTGGAGGAGGGGTG	1064
Db	134	GGAGATGTTTCCCGAGTATCCACTGGCTCCCGCAGCCCGCCAGTTTGGAGGAGGGGTG	193
QY	1065	CCGCAACAGCAGAGTCTCTGGACCTGACGAGGAGCGCGAGTTTGAACCCGGGAGCAG	1124
Db	194	CCGCAACAGCAGAGTCTCTGGACCTGACGAGGAGCGCGAGTTTGAACCCGGGAGCAG	253
QY	1125	AGCCAGTGGCCACCGTACCAATGGCATTCATGTACCGCGGGGTCTATGACTATCACT	1184
Db	254	AGCCAGTGGCCACCGTACCAATGGCATTCATGTACCGCGGGGTCTATGACTATCACT	313
QY	1185	GGCAACATCTACATCTACAATGGACCACTACTGGGGGGACACCGGGTCTCTGGAGACCTC	1244
Db	314	GGCAACATCTACATCTACAATGGACCACTACTGGGGGGACACCGGGTCTCTGGAGACCTC	373
QY	1245	CCAGCTACCCCGCAACCTCCATACCCCATTCGCCGAAGAGGGGACCCCTGGCCCTCCCGGG	1304
Db	374	CCAGCTACCCCGCAACCTCCATACCCCATTCGCCGAAGAGGGGACCCCTGGCCCTCCCGGG	433
QY	1305	CTCTCTACCCCGCAACCTCCATACCCCATTCGCCGAAGAGGGGACCCCTGGCCCTCCCGGG	1364
Db	434	CTCTCTACCCCGCAACCTCCATACCCCATTCGCCGAAGAGGGGACCCCTGGCCCTCCCGGG	493
QY	1365	GGTGGCCACACCTCTACAGGGGCCCCAAGAACCAATTTATCACCCTGACTGACTGAGT	1424
Db	494	GGTGGCCACACCTCTACAGGGGCCCCAAGAACCAATTTATCACCCTGACTGACTGAGT	553
QY	1425	CTGAGAAAGGCGAAGAGGGGGGACAAAGGGCACCTTCTCCCTTGGAGGCTGCCCTGCC	1484
Db	554	CTGAGAAAGGCGAAGAGGGGGGAC -AGGGCACCTTCTCCCTTGGAGGCTGCCCTGCC	612
QY	1485	CAGCTGGGATTACAGGGGCTCAGTAGGGCCCGGGGAAGCAGAGCCCTTAAGGATTAAAG	1544
Db	613	CAGCTGGGATTACAGGGGCTCAGTAGGGCCCGGGGAAGCAGAGCCCTTAAGGATTAAAG	672
QY	1545	GCTCAGACACCTCTGAGAGGAGGTGGCACTGGCTGGG--TACGGTCCCTCCACAGGAC	1602
Db	673	GCTCAGACACCTCTGAGAGGAGGTGGCACTGGCTGGGTTACGGTTGCCCTCCACAGGAC	732
QY	1603	TCCTCCCTACTGCC -TGAGCAAACTGTAG -GCCTCCCGGAGACCCACCCCTCCCTGGGG	1660

QY	287	GCTCTGGCAGCATCGAGCCCGAGCGGTGCTCCATATGCTCGGAGAACGACGACCTG	346
Db	240	GCTCTGGCAGCATCGAGCCCGAGCGGTGCTCCATATGCTCGGAGAACGACGACCTG	299
QY	347	CAGGACACAGGAAAGAAATCTATGAGCCCGAGCCCGAGCAGCTGTGCTCCGCTGCC	406
Db	300	CAGGACACAGGAAAGAAATCTATGAGCCCGAGCCCGAGCAGCTGTGCTCCGCTGCC	359
QY	407	GCAGGACCACTATGCTCAGCTAAATGTAGCCGATCCGCGACACAGTTGTGCCACATG	466
Db	360	GCAGGACCACTATGCTCAGCTAAATGTAGCCGATCCGCGACACAGTTGTGCCACATG	419
QY	467	TGCCGAGAATTCCTACAACGAGCAGCTGGAATACCTGACCATCTGCCAGCTGTGCCGCC	526
Db	420	TGCCGAGAATTCCTACAACGAGCAGCTGGAATACCTGACCATCTGCCAGCTGTGCCGCC	479
QY	527	CTGTGACCCAGTATGGGCTCGAGGAGATGGCCCTCGACAAAGCAAGCGAAGACCCA	586
Db	480	CTGTGACCCAGTATGGGCTCGAGGAGATGGCCCTCGACAAAGCAAGCGAAGACCCA	539
QY	587	GTGCCGCTGCGAGCGGATGTTCTGCTGCTGGCCCTCGAGTGTACACACTGCGA	646
Db	540	GTGCCGCTGCGAGCGGATGTTCTGCTGCTGGCCCTCGAGTGTACACACTGCGA	599
QY	647	GCTACTTCTGACTGCGCGCTGGCACTGAAGCGGAGCTCAAGATGAAGTTGGGAAGG	706
Db	600	GCTACTTCTGACTGCGCGCTGGCACTGAAGCGGAGCTCAAGATGAAGTTGGGAAGG	659
QY	707	TACAACCACTGCTGCTGCAAGGAGGCACTTCCAGATACCTCTCCCGCCAGCGC	766
Db	660	TACAACCACTGCTGCTGCAAGGAGGCACTTCCAGATACCTCTCCCGCCAGCGC	718
QY	767	CCGCTGCCAGCCCGAGGAGTGTGAGACCAAGGCTGCTGGAGGAGGAGGAGG	826
Db	719	CCGCTGCCAG -GCCACACCAAGGTTGAGAACCAAGGCTGCTGGAGGAGGAGGAGG	777
QY	827	TGCCAGTCCGACA -AAGCTGCAAAATCCATTAGAGCCACTGCCCCAGAGATGTGAGG	886
Db	778	TGCCAGTCCGACA -AAGCTGCAAAATCCATTAGAGCCACTGCCCCAGAGATGTGAGG	834
QY	887	ATC 889	
Db	835	AAC 837	

RESULT 11
BI821009
LOCUS 603035664F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5176935 5',
DEFINITION mRNA sequence.
ACCESSION BI821009
VERSION BI821009.1 GI:15932559
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 954)
NIH-MGC <http://imgc.ncl.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM1440 row: 0 column: 16
High quality sequence stop: 758.

```

Db 733 TCTCCCTACTGTAGCAACCTGAGTCCCTCCGGGAGACCAACCCACCCCTGGG 792
QY 1661 CTGCTCAGCCTCAGCAGCGGACGGCACATG-----ATACCAACTGCTGCCCA 1709
Db 793 GGCCTGCTCAGCCTCCAGGACGACGGACCGACATGATGATACCAAGCTGCTGCCCA 852
QY 1710 CTACAGACACGCGGACCGGACGACGGACCGA-GGGAGCCGCCACAC-GGTCACCTGCAA 1767
Db 853 CTAAGGCACGCGGACCGGACCGGACCGGACCGATGGAGCCGCCACACGCGGTCACTGCAA 912
QY 1768 GGAGTCACGGGCGCCCTTAAG 1790
Db 913 GGAGTCACGGGCGCCCTCTAAGG 935

RESULT 12
BO957322 895 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT_8777754 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6384424
5', mRNA sequence.
ACCESSION BO957322
VERSION BO957322.1 GI:22372800
KEYWORDS EST
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2579 row: g column: 17
High quality sequence stop: 645.
Location/Qualifiers
1..895
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6384424"
/clone_lib="NIH_MGC_40"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
Note: this is a NIH_MGC Library.
BASE COUNT 190 a 291 c 252 g 159 t 3 others
ORIGIN

Query Match 36.8%; Score 729; DB 14; Length 895;
Best Local Similarity 98.7%; Pred. No. le-154;
Matches 754; Conservative 0; Mismatches 8; Indels 2; Gaps 2;

QY 885 GGATCGCTGTCAAGAGCGGTCCGACGGAGAGGACCCCAATCTGTAGCTGGAAGCTGG 944
Db 132 GGATCGCTGTCAAGAGCGGTCCGACGGAGAGGAGGACCCCAATCTGTAGCTGGAAGCTGG 191
QY 945 GAGCCTCCGAAGGCCCATTCATCTCCCTGACTTGTGTACAGGCACATGCTACCCATTCT 1004
Db 192 GAGCCTCCGAAGGCCCATTCATCTCCCTGACTTGTGTACAGGCACATGCTACCCATTCT 251

```

```

QY 1005 GGAGATGTTTCCCCAGTATCCACTGGGTCCCGCAGCCCCCAGTTTGGAGCGAGGGTG 1064
Db 252 GGAGATGTTTCCCCAGTATCCACTGGGTCCCGCAGCCCCCAGTTTGGAGCGAGGGTG 311
QY 1065 CCGCAACAGCAGAGATCTCTGGACCTGACAGGAGGCGCAGTTGGAAACCCGGGAGAG 1124
Db 312 CCGCAACAGCAGAGATCTCTGGACCTGACAGGAGGCGCAGTTGGAAACCCGGGAGAG 371
QY 1125 AGCCAGGTGGCCCGCAGGTACCAATGGGATTCATGTCACCGCGGGTCTATGACTATCACT 1184
Db 372 AGCCAGGTGGCCCGCAGGTACCAATGGGATTCATGTCACCGCGGGTCTATGACTATCACT 431
QY 1185 GGCAACATCTACATCTACAATGGACCACTGTTGGGGGACCAACCGGGTCTCTGGAGACCTC 1244
Db 432 GGCAACATCTACATCTACAATGGACCACTGTTGGGGGACCAACCGGGTCTCTGGAGACCTC 491
QY 1245 CCAGCTACCCCGGAACTCCATACCCCATTCGCCAAGAGGGGACCTGGCCCTCCCGGG 1304
Db 492 CCAGCTACCCCGGAACTCCATACCCCATTCGCCAAGAGGGGACCTGGCCCTCCCGGG 551
QY 1305 CTCCTACACCCCGCAGGAGATGGCAAGGCTTGGCACCTAGCGAGAGACAGACACTGT 1364
Db 552 CTCCTACACCCCGCAGGAGATGGCAAGGCTTGGCACCTAGCGAGAGACAGACACTGT 611
QY 1365 GGTGCCACACCTCTAACAGGGGCCCAAGGAACCAATTTATCACCCATGACTGACTGAGT 1424
Db 612 GGTGCCACACCTCTAACAGGGGCCCAAGGAACCAATTTATCACCCATGACTGACTGAGT 671
QY 1425 CTGAGAAAAGCAGAAAGAGGGGGCAGACAGGACCTTCCCTTGGAGGTGCCCTGCC 1484
Db 672 CTGAGAAAAGCAGAAAGAGGGGGCAGACAGGACCTTCCCTTGGAGGTGCCCTGCC 731
QY 1485 CACGTGGATTTCACAGGGGCTGAGTAGGGCCGGGAAGCAGAGCCCTTAGGGATTAAAG 1544
Db 732 CACGTGGATTTCACAGGGGCTGAGTAGGGCCGGGAAGCAGAGCCCTTAGGGATTAAAG 790
QY 1545 GTCAGACACCTCTGAGAGCAGGTGGGCACCT-GGCTGGGTACGGTCCCTCCACAGGACT 1603
Db 791 GTCAGACACCTCTGAGAGCAGGTGGGCACCTGGGCTGGGTACGGTCCCTCCACAGGACT 850
QY 1604 CTCCTACTGCTGAGCAAACTGAGCGCTCCCGGACAGACCCAC 1647
Db 851 CTCCTACTGCTGAGCAAACTGAGCGCTCCCGGACAGCCANC 894

```

RESULT 13

BI911211

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BI911211 891 bp mRNA linear EST 16-OCT-2001
603062959F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5212047 5',
mRNA sequence.

BI911211

BI911211.1

GI:16174826

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 891)

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1AM1532 row: f column: 16

High quality sequence start: 26

High quality sequence stop: 849.

FEATURES

source

Location/Qualifiers
1. .891
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5212047"
/clone_lib="NIH_MGC_118"
/tissue_type="leukocyte"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH_MGC Library."
166 a 312 c 249 g 164 t

BASE COUNT	166 a	312 c	249 g	164 t
ORIGIN				
Query Match	36.5%; Score 723.2; DB 13; Length 891;			
Best Local Similarity	95.0%; Pred. No. 2.1e-153;			
Matches	781; Conservative	0; Mismatches	33; Indels	8; Gaps
3;				
Qy	86	GGGTGCACATCGGCGCTGAGTCCCTCCAGGCTCTGGGCTCGGCGAGCGCGCGCCACCG	145	
Db	30	GGGTGCACATCGGCGCTGAGTCCCTCCAGGCTCTGGGCTCGGCGAGCGCGCGCCACCG	89	
Qy	146	CTGCCAGAGCTGGGCGCTCTGCTTCCTCCAGGCGCCCGACCTGCTGGCGCCCTGG	205	
Db	90	CTGCCAGAGCTGGGCGCTCTGCTTCCTCCAGGCGCCCGACCTGCTGGCGCCCTGG	149	
Qy	206	CCGAGTGGCGCGATGCTTCCTGGGCTCGGCGAGCGCGCGCGCCCTGGCGCGCC	265	
Db	150	CCGAGTGGCGCGATGCTTCCTGGGCTCGGCGAGCGCGCGCGCCCTGGCGCGCC	209	
Qy	266	TCTGTGCTGGGCGCTCTTCGGGCTCTGGCGAGCTCGGCGAGCGCGCGCGCTCCCAT	325	
Db	210	TCTGTGCTGGGCGCTCTTCGGGCTCTGGCGAGCTCGGCGAGCGCGCGCTCCCAT	269	
Qy	326	TGCGTGGAGAACACGAGCTCGAGGACCGAGAAAGAAATCTATGAGCCCGACGACCG	385	
Db	270	TGCGTGGAGAACACGAGCTCGAGGACCGAGAAAGAAATCTATGAGCCCGACGACCG	329	
Qy	386	CATCTGCTGCTCCGCTCCCGCGAGGACCTATGCTCAGCTAACTAGCC-GCATCC	444	
Db	330	CATCTGCTGCTCCGCTCCCGCGAGGACCTATGCTCAGCTAACTAGCCGCTATCC	389	
Qy	445	GGGACACAGTTGTGCCACATGTGCCGAGAATTCCTACAAGGAGCAGCTGGAACCTGTA	504	
Db	390	GGGACACAGTTGTGCCACATGTGCCGAGAATTCCTACAAGGAGCAGCTGGAACCTGTA	449	
Qy	505	CCATCTGCCAGCTGTGCGCGCCCTGTGACCCAGTGTGGGCGCTCGAGGAGATGCGCCCT	564	
Db	450	CCATCTGCCAGCTGTGCGCGCCCTGTGACCCAGTGTGGGCGCTCGAGGAGATGCGCCCT	509	
Qy	565	GCACAAGCAAGGAGACCGAGTCCGCTGCGAGCGGGAATGTTCTGCTGCTGCTGG	624	
Db	510	GCACAAGCAAGGAGACCGAGTCCGCTGCGAGCGGGAATGTTCTGCTGCTGCTGG	569	
Qy	625	CCCTCGAGTGTACACATGCGAGCTACTTTCTGACTGCCCGCTGGCACTGAAGCCGAGC	684	
Db	570	CCCTCGAGTGTACACATGCGAGCTACTTTCTGACTGCCCGCTGGCACTGAAGCCGAGC	629	
Qy	685	TCAAAGATGAAGTTGGGAAGGTTAAACACCACTCGCTCCCTGCAAGGAGGCACTTCC	744	
Db	630	TCAAAGATGAAGTTGGGAAGGTTAAACACCACTCGCTCCCTGCAAGGAGGCACTTCC	689	
Qy	745	AGAATACCTCTCTCCCGAGCGCGCTGCGAGCCCGACACAGGTG-TGAGAACAAGGT	803	
Db	690	AGAATACCTCTCTCCCGAGCGCGCTGCGAGCCCGACACAGGTGTTGAGAACAAGGT	749	

Qy	804	CTGGTGGAGCGAGCTCCAG-----GCAGTCCCGAGTCCGACACAACTGCAAAATCCA	857	
Db	750	CTGGTGGAGCGAGCTCCAGGCGACTGCCCGAGTCCGACCAAACTGCAAAATCCA	809	
Qy	858	TTAGAGCCACTGCCCGGAGAGATGTCTAGATCGCTGCTCAAG	899	
Db	810	TTAGAGCAACTGGCGCCGAGAGATGTCTAGAGCAACCATGCTGATG	851	
RESULT 14				
BI818683				
LOCUS				
DEFINITION	603033289F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5174303 5',			
ACCESSION	BI818683			
VERSION	BI818683.1 GI:15930201			
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE				
AUTHORS	1 (bases 1 to 788)			
TITLE	NIH-MGC http://mgc.nci.nih.gov/.			
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)			
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: caapbs-remail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM11434 row: a column: 24 High quality sequence stop: 779.			
FEATURES				
source				
Location/Qualifiers				
1. 788				
/organism="Homo sapiens"				
/db_xref="taxon:9606"				
/clone="IMAGE:5174303"				
/clone_lib="NIH_MGC_115"				
/lab_host="DH10B"				
/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."				
BASE COUNT	146 a	287 c	223 g	132 t
ORIGIN				
Query Match	36.2%; Score 716.8; DB 13; Length 788;			
Best Local Similarity	96.8%; Pred. No. 5.6e-152;			
Matches	763; Conservative	0; Mismatches	22; Indels	3; Gaps
3;				
Qy	88	GTGCACATCGGCGCTGAGTCCCGTCCAGGCTCTGGGCTCGGGAGCGCGCCACCGC-	146	
Db	1	GTGCACATCGGCGCTGAGTCCCGTCCAGGCTCTGGGCTCGGGAGCGCGCCACCGC	60	
Qy	147	TGCCCCAGGACGTGCGGCGCTCTGCTTCTCCAGGCGCCCGACGTTGCTGGCGCGCTGGC	206	
Db	61	TGCCCCAGGACGTGCGGCGCTCTGCTTCTCCAGGCGCCCGACGTTGCTGGCGCGCTGGC	120	
Qy	207	CGAGTGGCGCGCATGCTCTGCTTGGGCGCACCTCTGCCCCCGGCGCTGGCGGGCGCT	266	
Db	121	CGAGTGGCGCGCATGCTCTGCTTGGGCGCACCTCTGCCCCCGGCGCTGGCGGGCGCT	180	
Qy	267	CTGGTGTGCTGGGCGCTCTTCTGGGCTCTGTGCGAGCATCGAGCCCGCGGCTGCTCCATAT	326	

```

Db 181 CTGTGCTGGGCTCTTCTGGGCTCTGTGCAGCATCGAGCCCGAGGGTGCCTCCATAT 240
Qy 327 GCGTCGGAGAACAGACTGTCAGGAGCAGGAGAAAGAAATCTATGAGCCCGCAGCCGC 386
Db 241 GCGTCGGAGAACAGACTGTCAGGAGCAGGAGAAAGAAATCTATGAGCCCGCAGCCGC 300
Qy 387 ATCTGCTGCTCCGCTGTCGCGCCCGCAGGACCTATCTCTAGCTAAATGTAGCCGATCCGG 446
Db 301 ATCTGCTGCTCCGCTGTCGCGCCCGCAGGACCTATCTCTAGCTAAATGTAGCCGATCCGG 360
Qy 447 GACACAGTTGTCGACATGTCGCGAGAAATCTCTACACGAGCAGCTGAACTACTGACC 506
Db 361 GACACAGTTGTCGACATGTCGCGAGAAATCTCTACACGAGCAGCTGAACTACTGACC 420
Qy 507 ATCTGCCAGCTGTGTCGCGCCCGCTGTGACCCAGTGTGCGCCTCGAGGAGATTGCCCCCTGC 566
Db 421 ATCTGCCAGCTGTGTCGCGCCCGCTGTGACCCAGTGTGCGCCTCGAGGAGATTGCCCCCTGC 480
Qy 567 ACAAGCAACGAGAGACCCAGTGCCTGCGAGCCGCGGAAATGTTCTGTGCTGGGCC 626
Db 481 ACAAGCAACGAGAGACCCAGTGCCTGCGAGCCGCGGAAATGTTCTGTGCTGGGCC 540
Qy 627 CTCGAGTGTACACACTGCGAGCTACTTCTGACTGCCGCTGCGACTGAGCCGAGCTC 686
Db 541 CTCGAGTGTACACACTGCGAGCTACTTCTGACTGCCGCTGCGACTGAGCCGAGCTC 600
Qy 687 AAAGATGAAGTTGGGAAGGTA-ACAACCACTGCTGCCCTGCAAGGCA-GGGCACTTCC 744
Db 601 AAAGATGAAGTTGGGAAGGTTAGCAACCACTGCTGCCCTGCAAGGCAAGGCGACTTCC 660
Qy 745 AGAATACCTCTCCCGCAGCGCCGCTGCGAGCCCGCAGCCAGAGTGTGAGAACCAAGTC 804
Db 661 AGAATACCTCTCTCCCGCAGCGCCGCTGCGAGCCCGCAGCCAGAGTGTGAGAACCAAGTC 720
Qy 805 TGGTGGAGGAGCTCCAGGCACTGCCAGTCCGACACAACTGCAAAATCCATTAGAGC 864
Db 721 TGGTGGAGGAGCTCCAGGCACTGCCAGTCCGACACAACTGCAAAATCCATTAGAGC 780
Qy 865 CACTGCC 872
Db 781 CATGCC 788

RESULT 15
Bi761249
LOCUS 603043491f1 NIH_MGC_116 Homo sapiens cDNA linear EST 25-SEP-2001
DEFINITION mRNA sequence.
ACCESSION Bi761249
VERSION Bi761249.1 GI:15752827
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 881)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11458 row: f column: 23
High quality sequence stop: 815.
Location/Qualifiers
1. .881

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5183638"
/lab_lib="NIH_MGC_116"
/lab_host="DH10B"
/notes="Organ: pooled colon, kidney, stomach; Vector:
pCMV-SpOrf6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH_MGC Library."
BASE COUNT 160 a 315 c 255 g 151 t
ORIGIN

```

```

Query Match 36.0%; Score 714.4; DB 13; Length 881;
Best Local Similarity 97.6%; Pred. No. 2e-151;
Matches 810; Conservative 0; Mismatches 11; Indels 9; Gaps 8;

Qy 60 GCCCGGCTGGCGCTCCCGGCTGGGTGCACATCGGCGCTGAGTCCGTCGCCAGGCT 119
Db 8 GCGCGGCTGGCGCTCCCGGCTGGGTGCACATCGGCGCTGAGTCCGTCGCCAGGCT 67
Qy 120 CTGGGCTCGGGCAGCC-GCCGCCACCGCTGCCAG-GACGTCGGGGCTCTCTGCTTCTC 177
Db 68 CTGGGCTCGGGCAGCTGCCGCCACCGCTGCCAGTGAAGTCGGGCTCTCTGCTTCTC 127
Qy 178 CCAGGCCCCCAGTGTGCTGGCGGCTGGCGAGTGGCGGCATCTCTGCTGCTTG-GGCC 236
Db 128 CCAGGCCCCCAGTGTGCTGGCGGCTGGCGAGTGGCGGCATCTCTGCTGCTTGCGCC 187
Qy 237 ACCTCTGCCCGCGGCTGGCTGGGGGCTCTGTGCTGGGCTCTTCTGCGCTCTGSCA 296
Db 188 ACCTCTGCCCGCGGCTGGCTGGGGGCTCTGTGCTGGGCTCTTCTGCGCTCTGSCA 247
Qy 297 GCATCGAGCCCCAGCGGTGCCCTCCATATGCGTCGGAGAACCAAGCTCGAGGACAG 356
Db 248 GCATCGAGCCCCAGCGGTGCCCTCCATATGCGTCGGAGAACCAAGCTCGAGGACAG 307
Qy 357 GAAAGGAATACTATGAGCCCCAGCACCAGCATCTGCTGCTCCCGCTGCCCGCAGGCA 416
Db 308 GAAAGGAATACTATGAGCCCCAGCACCAGCATCTGCTGCTCCCGCTGCCCGCAGGCA 367
Qy 417 TATGCTCTAGCTAAATGTAGCGCATCCGGACACAGTTGTGCCACATGTGCCGAGAAT 476
Db 368 TATGCTCTAGCTAAATGTAGCGCATCCGGACACAGTTGTGCCACATGTGCCGAGAAT 427
Qy 477 TCCTACAACGAGCAGTGGAACTACCTGACCATCTGCCAGCTGTGCCCGCTCTGTACCCA 536
Db 428 TCCTACAACGAGCAGTGGAACTACCTGACCATCTGCCAGCTGTGCCCGCTCTGTACCCA 487
Qy 537 GTGATGGGCTCGAGGAGATTGCCCCCTGCACAGCAAAAGGAGAGCCAGTGCCTGTC 596
Db 488 GTGATGGGCTCGAGGAGATTGCCCCCTGCACAGCAAAAGGAGAGCCAGTGCCTGTC 547
Qy 597 CAGCGGGAATGTTCTGCTGCTGGCGCTCGAGTGTACACACTCGAGCTACTTTCT 656
Db 548 CAGCGGGAATGTTCTGCTGCTGGCGCTCGAGTGTACACACTCGAGCTACTTTCT 607
Qy 657 GACTGCCCGCTGGCACTGAAGCCGAGCTCAAAGATGAAGTTGGAA-GGGTAACACCA 715
Db 608 GACTGCCCGCTGGCACTGAAGCCGAGCTCAAAGATGAAGTTGGAAAGCGGTAAACACCA 667
Qy 716 CTGGTCCCTCGAAGCAGGACACTTCCAGATACCTCTCCCGCAGCCCGCTGCTGCA 775
Db 668 CTGGTCCCTCGAAGCAGGACACTTCCAGATACCTCTCCCGCAGCCCGCTGCTGCA 727
Qy 776 GCCCCACACAGTGTGAGAACCC-AAGGCTGTGGTGGAG-GCAGCTCCAGGC-ACTGCCCA 832
Db 776 GCCCCACACAGTGTGAGAACCC-AAGGCTGTGGTGGAG-GCAGCTCCAGGC-ACTGCCCA

```

Db 728 GCGCCGCAACAGGTGTGAGAACCAAAAGTCTGGTGAGCGCAGCTCCAGGCAACTGCCCA 787
Qy 833 GTCCGACACAACTGCAAAA--TCCATTAGAGCCCACTGCCCCGAGAGAT 880
Db 788 GTCCGACACAACTGCAAAAAGTCCATTAGGAGCCCACTGCCCCGAGAGAT 837

Search completed: April 15, 2003, 16:26:18
Job time : 2303.32 secs

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:11:20 ; Search time 63.8651 Seconds
(without alignments)
9517.464 Million cell updates/sec

Title: US-09-917-372-2

Perfect score: 1982

Sequence: 1 gccccgcgcgcagctcgtc.....ggtatgggggaagatttgg 1982

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*

- 1: /cgn2.6/ptodata/1/ina/5A_COMB.seq:*
- 2: /cgn2.6/ptodata/1/ina/5B_COMB.seq:*
- 3: /cgn2.6/ptodata/1/ina/6A_COMB.seq:*
- 4: /cgn2.6/ptodata/1/ina/6B_COMB.seq:*
- 5: /cgn2.6/ptodata/1/ina/PCTUS_COMB.seq:*
- 6: /cgn2.6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	72	3.6	1557	US-08-385-229-3	Sequence 3, Appli
2	72	3.6	1641	US-08-385-229-1	Sequence 1, Appli
3	72	3.6	1641	US-08-650-000-1	Sequence 1, Appli
4	72	3.6	1641	5395760-1	Patent No. 5395760
5	72	3.6	2224	US-08-477-347-2	Sequence 2, Appli
6	72	3.6	2224	US-08-476-862-1	Sequence 1, Appli
7	72	3.6	3683	US-09-844-634-3	Sequence 3, Appli
8	62.6	3.2	691	US-08-266-080B-12	Sequence 12, Appl
9	62.6	3.2	691	PCT-US95-05423-12	Sequence 12, Appl
10	62.6	3.2	705	US-09-326-394-3	Sequence 3, Appli
11	62.6	3.2	705	US-09-580-235-1	Sequence 1, Appli
12	62.6	3.2	705	US-09-580-235-3	Sequence 3, Appli
13	62.6	3.2	705	US-09-580-235-5	Sequence 5, Appli
14	62.6	3.2	705	US-09-580-235-7	Sequence 7, Appli
15	62.6	3.2	705	US-09-580-181-1	Sequence 1, Appli
16	62.6	3.2	705	US-09-580-181-3	Sequence 3, Appli
17	62.6	3.2	705	US-09-580-181-5	Sequence 5, Appli
18	62.6	3.2	705	US-09-580-181-7	Sequence 7, Appli
19	62.6	3.2	705	US-09-102-530-1	Sequence 1, Appli
20	62.6	3.2	705	US-09-102-530-3	Sequence 3, Appli
21	62.6	3.2	705	US-09-102-530-5	Sequence 5, Appli
22	62.6	3.2	705	US-09-102-530-7	Sequence 7, Appli
23	62.4	3.1	579	US-09-146-950-3	Sequence 3, Appli
24	62.4	3.1	591	US-09-146-950-19	Sequence 19, Appl
25	62.4	3.1	1596	US-09-146-950-17	Sequence 17, Appl
26	62.4	3.1	1724	US-08-509-024-1	Sequence 1, Appli
27	62.4	3.1	1724	US-09-333-279-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-08-385-229-3

; Sequence 3, Application US/08385229

; Patent No. 5605690

; GENERAL INFORMATION:

; APPLICANT: Jacobs, Cindy A.

; APPLICANT: Smith, Craig A.

; TITLE OF INVENTION: Method of Treating TNF-Dependent

; TITLE OF INVENTION: Inflammation Using Tumor Necrosis Factor Antagonists

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Immunex Corporation

; STREET: 51 University Street

; CITY: Seattle

; STATE: Washington

; COUNTRY: U.S.A.

; ZIP: 98101

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/385,229

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/07/946,236

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Wright, Christopher L.

; REGISTRATION NUMBER: 31,680

; REFERENCE/DOCKET NUMBER: 2503

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 587-0430

; TELEFAX: (206) 587-0606

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1557 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; IMMEDIATE SOURCE:

; CLONE: TNFR/FC Fusion Protein

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..1557

Sequence 1, Appli
Sequence 1, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 4, Appli

```
;
;
; NAME/KEY: mat_peptide
; LOCATION: 1..1554
US-08-385-229-3

Query Match 3.6%; Score 72; DB 1; Length 1557;
Best Local Similarity 51.3%; Pred. No. 8e-08;
Matches 193; Conservative 0; Mismatches 180; Indels 3; Gaps 1;

QY 241 CTGCCCCGGGCTGGCTGGGGGCTCTGGTGGTGGGGCTCTTGGGGCTCTTGGGGCTCTTGGCAGCAT 300
DB 95 CCGTGGCGCTGGGGCGCGCTGGCGCTGGAGCTGGAGCTCTGGGCTGGCGCGCAGCGCT 154
QY 301 CGAGCCCCAGGGCTGCTCCATATGCTGGGAGAACACGACCTGCGAGGACGAGGAA 360
DB 155 TGCCCGCCAGGTGGCATTTACACCTAGCCCGCGGAGCGAGCATGCGGGCTCA 214
QY 361 AGGAATATATGAGCCCCAGCAGCATCTGCTGCTGCCGCTGCCGCGCAGGACCATATG 420
DB 215 GAGAATCTTCTGTACCAAGACCTGCGGACCGGTGTGTGACTCTCTGTGAGGACGACAT 334
QY 421 TCTAGCTAAATGTAGCGCATCGGGACACAGTTTGTGCCACATGTGCCGAGATTCCT 480
DB 275 CAAGAATCTTCTGTACCAAGACCTGCGGACCGGTGTGTGACTCTCTGTGAGGACGACAT 334
QY 481 ACAAGAGCTAAATGTAGCGCATCGGGACACAGTTTGTGCCACATGTGCCGAGATTCCT 480
DB 335 ACACCCAGCTCTGGAACCTGGGTTCGCCAGTGTCTGAGCTGTGGTCCCG---CTGTAGCT 391
QY 541 TGGGCTCTGAGGAGATTGCCCGCTGCAACAAGCAAGCAAGCAAGCAAGCAAGCAAG 600
DB 392 CTGACCAAGTGGAACTCAAGCTGCACTCGGAGAACAGACCGCATCTGTCCACCTGCAGGC 451

QY 601 CGGGAATCTTCTGTGC 616
DB 452 CCGGCTGTACTGCGC 467

RESULT 2
US-08-385-229-1
; Sequence 1, Application US/08385229
; Patent No. 5605690
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Cindy A.
; APPLICANT: Smith, Craig A.
; TITLE OF INVENTION: Method of Treating TNF-Dependent
; TITLE OF INVENTION: Inflammation Using Tumor Necrosis Factor Antagonists
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESS: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/385,229
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/946,236
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wight, Christopher L.
; REGISTRATION NUMBER: 31,680
; REFERENCE/DOCKET NUMBER: 2503
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
```

```
;
;
; TELEFAX: (206) 587-0606
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1641 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; CELL TYPE: Fibroblast
; CELL LINE: WI-26 VA4
; IMMEDIATE SOURCE:
; LIBRARY: WI-26 VA4
; CLONE: Clone 1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 88..1473
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 154..1470
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 88..153
US-08-385-229-1

Query Match 3.6%; Score 72; DB 1; Length 1641;
Best Local Similarity 51.3%; Pred. No. 8.1e-08;
Matches 193; Conservative 0; Mismatches 180; Indels 3; Gaps 1;

QY 241 CTGCCCCGGGCTGGCTGGGGGCTCTGGTGGTGGGGCTCTTGGGGCTCTTGGGGCTCTTGGCAGCAT 300
DB 95 CCGTGGCGCTGGGGCGCGCTGGCGCTGGAGCTGGAGCTCTGGGCTGGCGCGCAGCGCT 154
QY 301 CGAGCCCCAGGGCTGCTCCATATGCTGGGAGAACACGACCTGCGAGGACGAGGAA 360
DB 155 TGCCCGCCAGGTGGCATTTACACCTAGCCCGCGGAGCGAGCATGCGGGCTCA 214
QY 361 AGGAATATATGAGCCCCAGCAGCATCTGCTGCTGCCGCTGCCGCGCAGGACCATATG 420
DB 215 GAGAATCTTCTGTACCAAGACCTGCGGACCGGTGTGTGACTCTCTGTGAGGACGACAT 334
QY 421 TCTAGCTAAATGTAGCGCATCGGGACACAGTTTGTGCCACATGTGCCGAGATTCCT 480
DB 275 CAAGAATCTTCTGTACCAAGACCTGCGGACCGGTGTGTGACTCTCTGTGAGGACGACAT 334
QY 481 ACAAGAGCTAAATGTAGCGCATCGGGACACAGTTTGTGCCACATGTGCCGAGATTCCT 480
DB 335 ACACCCAGCTCTGGAACCTGGGTTCGCCAGTGTCTGAGCTGTGGTCCCG---CTGTAGCT 391
QY 541 TGGGCTCTGAGGAGATTGCCCGCTGCAACAAGCAAGCAAGCAAGCAAGCAAGCAAG 600
DB 392 CTGACCAAGTGGAACTCAAGCTGCACTCGGAGAACAGACCGCATCTGTCCACCTGCAGGC 451

QY 601 CGGGAATCTTCTGTGC 616
DB 452 CCGGCTGTACTGCGC 467

RESULT 3
US-08-650-000-1
; Sequence 1, Application US/08650000
; Patent No. 5945397
; GENERAL INFORMATION:
; APPLICANT: Smith, Craig A.
; APPLICANT: Goodwin, Raymond G.
; APPLICANT: Beckmann, M. Patricia
; TITLE OF INVENTION: Tumor Necrosis Factor Receptors
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
```



```

; FILING DATE: 09-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/547,750
; FILING DATE: 02-JULY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/505,441
; FILING DATE: 06-APRIL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER: SYNE200C5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 691 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-266-080B-12

Query Match
Best Local Similarity 3.2%; Score 62.6; DB 1; Length 691;
Matches 165; Conservative 0; Mismatches 149; Indels 3; Gaps 1;

QY 300 TGCAGCCCGCCAGCGGTGCTCCATATGCTGCGGAGAACACAGACCTGCGAGGACCCAGGAA 359
Db 139 TTGCCCGCCCGCCAGGTGCGCATTTACACCTACGCTCCGGAACCGGGTTCTACCTGCCGGCTC 198
QY 360 AAGGAATACATGAGCCCGCCAGCAGCCGATCTGCTGCTCCCGTTCGCCCGCCAGGACCTAT 419
Db 199 AGAGAATACATGAGCAGACAGCTCAGATGCTGTCGAGCAAGTGTCTCGCGGGCCCAACAT 258
QY 420 GTCTCAGCTAAATGTAGCGCATCCGCGGACACAGTTTGTGCCACATGTGCCGAGATTC 479
Db 259 GCAAAAGTCTTCTGTACCAAGACCTCGGACACCGGTGTGTGACTCTCTGTGAGGACAGCACA 318
QY 480 TACAAGCAGCAGTGAACATACCTGACCATCTGCCAGCTGTGCCGCCCTGTGACCCAGTG 539
Db 319 TACACCCAGCTCTGGAACCTGGGTTCGCCAGTGTGAGTGTGGCTCCCG---CTGTAGC 375
QY 540 ATGGCCCTCGAGGAGATTGCCCCCTGTCACAGCAACGAGAACGCCAGTGTGGCGGTGCCAG 599
Db 376 TCTGACCAAGTGGAAACTCAAGCCTGCACTCGGGAACAGAACCGCATCTGCACCTGCAGG 435
QY 600 CCGGGAATGTTCTGTGC 616
Db 436 CCGGCTGTACTGCGC 452

RESULT 9
PCT-US95-05423-12
; Sequence 12, Application PC/TUS9505423
; GENERAL INFORMATION:
; APPLICANT: Jack Lile
; APPLICANT: Tadahiko Kohno
; APPLICANT: Duane Bonam
; APPLICANT: Mary S. Rosendahl
; TITLE OF INVENTION: Production of Biologically Active
; TITLE OF INVENTION: Recombinant Neurotrophic Protein
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson & Bratschun, L.L.C.
; STREET: 8400 E. Prentice Avenue, Suite 200
; CITY: Englewood
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MG storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
```

```

; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/05423
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/266,090
; FILING DATE: 27-JUNE-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/240,122
; FILING DATE: 09-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/087,912
; FILING DATE: 06-JULY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/680,681
; FILING DATE: 04-APRIL-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/594,126
; FILING DATE: 09-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/547,750
; FILING DATE: 02-JULY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/505,441
; FILING DATE: 06-APRIL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER: SYNE200/PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 691 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
PCT-US95-05423-12

Query Match
Best Local Similarity 3.2%; Score 62.6; DB 5; Length 691;
Matches 165; Conservative 0; Mismatches 149; Indels 3; Gaps 1;

QY 300 TGCAGCCCGCCAGCGGTGCTCCATATGCTGCGGAGAACACAGACCTGCGAGGACCCAGGAA 359
Db 139 TTGCCCGCCCGCCAGGTGCGCATTTACACCTACGCTCCGGAACCGGGTTCTACCTGCCGGCTC 198
QY 360 AAGGAATACATGAGCCCGCCAGCAGCCGATCTGCTGCTCCCGTTCGCCCGCCAGGACCTAT 419
Db 199 AGAGAATACATGAGCAGACAGCTCAGATGCTGTCGAGCAAGTGTCTCGCGGGCCCAACAT 258
QY 420 GTCTCAGCTAAATGTAGCGCATCCGCGGACACAGTTTGTGCCACATGTGCCGAGATTC 479
Db 259 GCAAAAGTCTTCTGTACCAAGACCTCGGACACCGGTGTGTGACTCTCTGTGAGGACAGCACA 318
QY 480 TACAAGCAGCAGTGAACATACCTGACCATCTGCCAGCTGTGCCGCCCTGTGACCCAGTG 539
Db 319 TACACCCAGCTCTGGAACCTGGGTTCGCCAGTGTGAGTGTGGCTCCCG---CTGTAGC 375
QY 540 ATGGCCCTCGAGGAGATTGCCCCCTGTCACAGCAACGAGAACGCCAGTGTGGCGGTGCCAG 599
Db 376 TCTGACCAAGTGGAAACTCAAGCCTGCACTCGGGAACAGAACCGCATCTGCACCTGCAGG 435
QY 600 CCGGGAATGTTCTGTGC 616
Db 436 CCGGCTGTACTGCGC 452

RESULT 10
US-09-326-394-3
; Sequence 3, Application US/09326394
; Patent No. 6306820
```


420 GTCTCAGCTAAATGTAGCGGCATCGGGGACACAGTTTGTGCCACATGTGCCGAGAATTC 479

QY 360 AAGGAATACTATGAGCCC

Db 1 TTGCCGCGCCAGGTGCATTACACCTACGCCCGGGAGCCCCGGAGCACATCGCGCTC 60

QY 360 AAGGAATACTATGAGCCCCAGCACGCATCTGCTCTCCGCTGCCCGCCAGGCACTTAT 419

Db 61 AGAGATACTATGACGACACCTCAGATGTCTGCAGCAATGCTCGCCGGGGCAACAT 120
QY 420 GTCTCAGCTAAATGAGCGCATCCGGGACACAGTTTGTGCCACATGTGCCGAGAAATCC 479
Db 121 GCAAAAGTCTTGTACCAAGACCTCGGACACCGTGTGTGACTCTCTGTGAGGACAGCACA 180
QY 480 TACACGAGCAGTGAACCTGACCTGACCATCTGCCAGCTGTGCCGCCCTGTGACCCAGTG 539
Db 181 TACACCGAGCTGTGAACTGGGTTCCTCGAGTGTGAGCTGTGGTCCCG---CTGTAGC 237
QY 540 ATGGGCTCGAGGAGATTGCCCTTGCACAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 599
Db 238 TCTGACCGAGTGAACCTGAACTGCACTCGGACAGAACCGCATCTGACCTGCAGG 297
QY 600 CCGGGAATGTTCTGTGC 616
Db 298 CCCGGCTGTACTGCGC 314

RESULT 14
US-09-580-235-7
; Sequence 7, Application US/09580235
; Patent No. 6433158
; GENERAL INFORMATION:
; APPLICANT: Pettit, Dean
; TITLE OF INVENTION: Site Specific Protein Modification
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Janis C Henry
; STREET: 51 University
; CITY: Seattle
; STATE: WA
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/580,235
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/102,530
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Henry, Janis C
; REGISTRATION NUMBER: 34,347
; REFERENCE/DOCKET NUMBER: 2637
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)470-4189
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 705 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..705
US-09-580-235-7

Query Match 3.2%; Score 62.6; DB 4; Length 705;
Best Local Similarity 52.1%; Pred. No. 9.8e-06;
Matches 165; Conservative 0; Mismatches 149; Indels 3; Gaps 1;

QY 300 TCGAGAGCCCGGGTGCCTCCATATGCTGGGAGAACCAAGCAAGCAAGCAAGCAAGCAAG 359

Db 1 TTGCCCCGCCAGGTGGCATTTACCCCTACGCCCTAGCCCCGGAGACCCGGGAGACATGCGCGGTC 60
QY 360 AAGGAATACTATGAGCCCCAGCAGCAGCTATGCTGTCTCCCGTGCCTCCGCGGACGACCTAT 419
Db 61 AGAGATACTATGAGCAGACAGCTCAGATGTCTGCAGCAATGCTCGCGGGCCAAAT 120
QY 420 GTCTCAGCTAAATGAGCGCATCCGGGACACAGTTTGTGCCACATGTGCCGAGAAATCC 479
Db 121 GCAAAAGTCTTGTACCAAGACCTCGGACACCGTGTGACTCTCTGTGAGGACAGCACA 180
QY 480 TACACGAGCAGTGAACCTGACCTGACCATCTGCCAGCTGTGCCGCCCTGTGACCCAGTG 539
Db 181 TACACCGAGCTGTGAACTGGGTTCCTCGAGTGTGAGCTGTGGTCCCG---CTGTAGC 237
QY 540 ATGGGCTCGAGGAGATTGCCCTTGCACAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 599
Db 238 TCTGACCGAGTGAACCTGAACTGCACTCGGACAGAACCGCATCTGACCTGCAGG 297
QY 600 CCGGGAATGTTCTGTGC 616
Db 298 CCCGGCTGTACTGCGC 314

RESULT 15
US-09-580-181-1
; Sequence 1, Application US/09580181
; Patent No. 6441136
; GENERAL INFORMATION:
; APPLICANT: Pettit, Dean
; TITLE OF INVENTION: Site Specific Protein Modification
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Janis C Henry
; STREET: 51 University
; CITY: Seattle
; STATE: WA
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/580,181
; FILING DATE: 26-MAY-2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/102,530
; FILING DATE: 22-JUN-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Henry, Janis C
; REGISTRATION NUMBER: 34,347
; REFERENCE/DOCKET NUMBER: 2637
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)470-4189
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 705 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..705
US-09-580-181-1

Query Match 3.2%; Score 62.6; DB 4; Length 705;
Best Local Similarity 52.1%; Pred. No. 9.8e-06;

Matches	165;	Conservative	0;	Mismatches	149;	Indels	3;	Gaps	1;
QY	300	TCGCAGCCCCCAGGGGTGCTCCATATGCGTTCGAGAAACAGACACTCGACGGAGACCAGAA	359						
Dd	1	TTGCCCGCCCAGGTGGCATTTACACCTACGCCCCGGAGCCCGGGAGCACATCGCGCTC	60						
QY	360	AAGGAATACTATAGCCCCCAGCACCGCATGTGTGCTCCCGTGCCCGCCAGGCACCTAT	419						
Dd	61	AGAATACTATTACCCAGACAGCTCAGATGCTGTGCAGCAAAATGCTCGCCGGGCCAACAT	120						
QY	420	GTCTCAGCTAAATGTAGCCGCATCCGGGACACAGTTTGTGCCACATGTGGCGAATTC	479						
Dd	121	GCAAAGTCTTCTGTACCAAGACCTCGGACACCGTGTGTACTCTGTGAGGACAGACA	180						
QY	480	TACAACGAGCACTGGAATACTACCTGACCATGTGCCAGCTGTGCCGCCCTGTGACCCAGTG	539						
Dd	181	TACACCAGCTCTGGAATCTGGTTCGCCAGTGTCTTGAGCTGTGGCTCCCG---CTGTAGC	237						
QY	540	ATGGGCTTCGAGGAGATTGCCCTGTGCACGAAGCAAAACGGAGACCCAGTCCGCTGCCAG	599						
Dd	238	TCTGACCAGGTGGAATACTCAAGCCTGCACTCGGGAACAGAACCCATCTGCACCTGCAGG	297						
QY	600	CCGGGAATGTTCTGTGC	616						
Dd	298	CCGGCTGGTACTGCGC	314						

Search completed: April 15, 2003, 16:28:26
Job time : 71.8651 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:04:15 ; Search time 1210.24 seconds
(without alignments)
15606.600 Million cell updates/sec

Title: US-09-917-372-2_COPY_629_1277

Perfect score: 649

Sequence: 1 cgagtggtacactcgagc.....aacctccatacccattccc 649

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg_hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htg_mus.*
- 34: em.htg_pln.*
- 35: em.htg_rod.*
- 36: em.htg_mam.*
- 37: em.htg_vrt.*
- 38: em.sv.*
- 39: em.htgo_hum.*
- 40: em.htgo_mus.*
- 41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	531	81.8	2091	9	AK095418	AK095418 Homo sapi
2	531	81.8	2136	6	AX331906	AX331906 Sequence
3	531	81.8	2136	6	AX332212	AX332212 Sequence
4	531	81.8	2136	6	AX409488	AX409488 Sequence
5	531	81.8	2136	9	HUMTNFRP	L04270 Homo sapien
6	529.4	81.6	2148	9	AK027080	AK027080 Homo sapi
7	529.4	81.6	2161	9	BC026262	BC026262 Homo sapi
8	270	41.6	1614	10	MUSLYNPHOB	L38423 Mus musculus
9	270	41.6	2076	10	MMU29173	U29173 Mus musculus
c 10	235.4	36.3	140026	9	AC005840	AC005840 Homo sapi
c 11	198	30.5	1605	9	HUMTUMNEC	L04489 Homo sapien
c 12	117.6	18.1	178228	2	AC128082	AC128082 Rattus no
c 13	99	15.3	323	9	HS27B9R	Z60529 H.sapiens C
c 14	97.6	15.0	193	9	HS27B9F	Z60528 H.sapiens C
c 15	82.4	12.7	187998	2	AC125909	AC125909 Rattus no
c 16	72	11.1	187998	2	AC125909	AC125909 Rattus no
c 17	66.8	10.3	2048	10	MMU30798	U30798 Mus musculus
c 18	45	6.9	125020	9	AF429315	AF429315 Homo sapi
c 19	41.6	6.4	125020	9	AF429315	AF429315 Homo sapi
c 20	41.4	6.4	107109	9	HSB781B1	AL118522 Human DNA
c 21	41	6.3	58311	2	AC110596	AC110596 Homo sapi
c 22	40.8	6.3	837	6	AX027018	AX027018 Sequence
c 23	40.8	6.3	845	6	AX027450	AX027450 Sequence
c 24	40.8	6.3	1276	4	AF248545	AF248545 Sus scrof
c 25	40	6.2	134092	2	H0702G05	AL442105 Oryza sat
c 26	39.8	6.1	167303	9	AP001527	AP001527 Homo sapi
c 27	39.8	6.1	189637	2	AC009765	AC009765 Homo sapi
c 28	39.8	6.1	199891	9	CNSO080	AL079303 Homo chr
c 29	39.4	6.1	91927	9	AC004771	AC004771 Homo sapi
c 30	39.2	6.0	519	6	AR076920	AR076920 Sequence
c 31	39.2	6.0	519	6	AR078310	AR078310 Sequence
c 32	39.2	6.0	519	6	AR085413	AR085413 Sequence
c 33	39.2	6.0	519	6	AR103369	AR103369 Sequence
c 34	39.2	6.0	519	6	AR169226	AR169226 Sequence
c 35	39.2	6.0	834	6	AR1027007	AR1027007 Sequence
c 36	39.2	6.0	1004	6	AR138075	AR138075 Sequence
c 37	39.2	6.0	1004	6	AR153584	AR153584 Sequence
c 38	39.2	6.0	1004	6	I07284	I07284 Sequence 31
c 39	39.2	6.0	1004	9	HSCDW40	X60592 Human CDW40
c 40	39.2	6.0	1157	9	BC012419	BC012419 Homo sapi
c 41	39.2	6.0	154011	2	AC018571	AC018571 Homo sapi
c 42	39.2	6.0	159983	2	AC040958	AC040958 Homo sapi
c 43	39.2	6.0	191583	9	AC015651	AC015651 Homo sapi
c 44	39.2	6.0	203613	2	AC118575	AC118575 Lemur cat
c 45	39	6.0	73565	2	AC121318	AC121318 Mus muscu

ALIGNMENTS

RESULT 1
AK095418
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AK095418
Homo sapiens CDNA FLJ138099 fis, clone D3OST1000238, highly similar
to LYMPHOTOXIN-BETA RECEPTOR PRECURSOR.
AK095418
AK095418.1 GI:21754669
oligo capping; fis (full insert sequence).
Homo sapiens CD34+ Cells cDNA to mRNA, clone_lib:D3OST1
clone:D3OST1000238.
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Oshima,A., Takahashi-Fujii,A., Tanase,T., Imose,N., Takeuchi,K.,

2091 bp mRNA linear PRI 15-JUL-2002

TITLE	NEDO human cdNA sequencing project
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 2091)
AUTHORS	Isogai, T. and Yamamoto, J.
TITLE	Direct Submission
JOURNAL	Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7 Kazusa-Kamatori, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT	NEDO human cdNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cdNA full insert sequencing: Research Association for Biotechnology (RAB); cdNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

[illegible]

Db	1046	TTCTGGAGATGTTTCCCCAGTATCCACTGGGCTCCCGCAGCCCAAGTCTTTGGAGGCAGG	1105
Qy	433	GGTCCGCAACAGCAGAGTCTCTGGAGCTGACACAGGAGCGCAGTTGGAACCCGGGA	492
Db	1106	GGTCCGCAACAGCAGAGTCTCTGGACCTGACACAGGAGCGCAGTTGGAACCCGGGA	1165
Qy	493	CGAGAGCAGGTGCCACCGGTACCAATGGCAATTCATGTACCGGCGGGTCTATGACTAT	552
Db	1166	CGAGAGCAGGTGCCACCGGTACCAATGGCAATTCATGTACCGGCGGGTCTATGACTAT	1225
Qy	553	CAC'TGGCAACATCTACATCTACATGACACAGTACTGGGGGACCAACCGGTCTTGAGA	612
Db	1226	CAC'TGGCAACATCTACATCTACATGACACAGTACTGGGGGACCAACCGGTCTTGAGA	1285
Qy	613	CCTCCAGTACCCCGGAACCTCCATACCCCAATTC	649
Db	1286	CCTCCAGTACCCCGGAACCTCCATACCCCAATTC	1322
RESULT 2			
AX331906			
LOCUS	AX331906	2136 bp	DNA linear PAT 09-JAN-2002
DEFINITION	Sequence 2415 from Patent WO0194629.		
ACCESSION	AX331906		
VERSION	AX331906.1	GI:18122540	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 Young, P. E., Augustus M., Carter, K. C., Ebner, R., Endress, G.,		
TITLE	Horrigan, S., Soppet, D. R. and Weaver, Z.		
JOURNAL	Cancer gene determination and therapeutic screening using signature		
FEATURES	gene sets		
source	Patent: WO 0194629-A 2415 13-DEC-2001;		
	Avalon Pharmaceuticals (US)		
	Location/Qualifiers		
	1..2136		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	446 a	706 c	608 g
ORIGIN	376 t		
Query Match	81.8%;	Score 531;	DB 6; Length 2136;
Best Local Similarity	85.7%;	Pred. No. 2.6e-129;	
Matches 649;	Conservative 0;	Mismatches 0;	Indels 108; Gaps 1;
Qy	1	CGAGTGTACACATCGCAGGTACTTCTGACTGCCGCCCTGGCACTGAAGCCGAGCTCAA	60
Db	579	CGAGTGTACACATCGCAGGTACTTCTGACTGCCGCCCTGGCACTGAAGCCGAGCTCAA	638
Qy	61	AGATGAAGTTGGGAAGGTTAAACACACTGGTCCCTGCAAGCAGGCACTTCCAGAA	120
Db	639	AGATGAAGTTGGGAAGGTTAAACACACTGGTCCCTGCAAGCAGGCACTTCCAGAA	698
Qy	121	TACCTCTCCCGAGCGCCGCTGCCAGCCCCACACAGTGTGAGAACCAAGTCTGGT	180
Db	699	TACCTCTCCCGAGCGCCGCTGCCAGCCCCACACAGTGTGAGAACCAAGTCTGGT	758
Qy	181	GGAGGCAGCTCCAGGCAGTGCCTCCAGTCCGACACAACTGCAAAAATCCATTAGAGCCACT	240
Db	759	GGAGGCAGCTCCAGGCAGTGCCTCCAGTCCGACACAACTGCAAAAATCCATTAGAGCCACT	818
Qy	241	GCCCCAGAGATGTC	256
Db	819	GCCCCAGAGATGTCAGGAACCATGCTGATGCTGGCCGTTCTGCTGCCACTGGCCTTCTT	878
Qy	257	-----	256
Db	879	TCTGCTCCTTGGCACCGCTTCTCTCTGCATCTGGAAGACCAACCTTCTCTCTGCAGAA	938
Qy	257	----GGATCGCTGCTCAAGAGCGCTCCGAGGGAGAGGGACCCAAATCCTCTAGCTGGAAG	312

```

Db 939 ACTGGATCGCTCTCAAGAGCGTCCGAGGAGGAGGACCAATCCTGTAGTGAAG 998
      |||
QY 313 CTGGAGCCTCCGAGGCCATCCATCTCCTGACTTGGTACAGCCACTGCTACCCAT 372
      |||
Db 999 CTGGAGCCTCCGAGGCCATCCATCTCCTGACTTGGTACAGCCACTGCTACCCAT 1058
      |||
QY 373 TTCTGAGATGTTTCCCAAGTATCCACTGGGTCCCGCAGGCCCAAGTTTGGAGGCAGG 432
      |||
Db 1059 TTCTGAGATGTTTCCCAAGTATCCACTGGGTCCCGCAGGCCCAAGTTTGGAGGCAGG 1118
      |||
QY 433 GTTCCGCGAACAGCAGAGTCTCTGTGACCTGACAGGAGGCCGAGTTTGAACCCGGGGA 492
      |||
Db 1119 GTTCCGCGAACAGCAGAGTCTCTGTGACCTGACAGGAGGCCGAGTTTGAACCCGGGGA 1178
      |||
QY 493 GCAGAGCCAGGTGGCCACGCTACCAATGCGATTCATGTCACCGCGGCTCTATGACTAT 552
      |||
Db 1179 GCAGAGCCAGGTGGCCACGCTACCAATGCGATTCATGTCACCGCGGCTCTATGACTAT 1238
      |||
QY 553 CACTGGCAACATCTACATCTACAATGGACAGTACTGGGGGAGCACACCGGGTCTTGAGA 612
      |||
Db 1239 CACTGGCAACATCTACATCTACAATGGACAGTACTGGGGGAGCACACCGGGTCTTGAGA 1298
      |||
QY 613 CCTCCAGCTACCCCGAACCTCCATACCCCATTTCCC 649
      |||
Db 1299 CCTCCAGCTACCCCGAACCTCCATACCCCATTTCCC 1335
      |||

RESULT 3
AX332212 AX332212 2136 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 2721 from Patent WO0194629.
DEFINITION AX332212
ACCESSION AX332212
VERSION AX332212.1 GI:18122846
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horrigan,S., Soppet,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 2721 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
source
Location/Qualifiers
i. .2136
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 446 a 706 c 608 g 376 t
ORIGIN
Query Match 81.8%; Score 531; DB 6; Length 2136;
Best Local Similarity 85.7%; Pred. No. 2.6e-129;
Matches 649; Conservative 0; Mismatches 0; Indels 108; Gaps 1;

QY 1 CGAGTGTACACTGCGAGCTACTTTCTGACTGCCCGCTGGCAGCTGAAGCCGAGTCAA 60
      |||
Db 579 CGAGTGTACACTGCGAGCTACTTTCTGACTGCCCGCTGGCAGCTGAAGCCGAGTCAA 638
      |||
QY 61 AGATGAAGTTGGGAAGGGTAACAACCACTGCGTCCCTCCAGGAGGAGGCACTTCCAGAA 120
      |||
Db 639 AGATGAAGTTGGGAAGGGTAACAACCACTGCGTCCCTCCAGGAGGAGGCACTTCCAGAA 698
      |||
QY 121 TACCTCCTCCCGAGCGCCGCTGCCAGCCGCCACACAGGTGTGAGAACCAAGTCTGTT 180
      |||
Db 699 TACCTCCTCCCGAGCGCCGCTGCCAGCCGCCACACAGGTGTGAGAACCAAGTCTGTT 758
      |||
QY 181 GGAGGAGCTCCAGGCACTGCCAGTCCGACACAACTCGAAGAAATTCATTTAGAGCCACT 240
      |||
Db 759 GGAGGAGCTCCAGGCACTGCCAGTCCGACACAACTCGAAGAAATTCATTTAGAGCCACT 818
      |||

```

```

QY 241 GCCCCAGAGATGTCA----- 256
      |||
Db 819 GCCCCAGAGATGTCAAGAACCATGCTGATGCTGGCGTTCTGCTGCCACTGGCCTTCTT 878
      |||
QY 257 ----- 256
      |||
Db 879 TCTGCTCCTTGGCACCGTCTTCTCCTGCACTCTGGAAGAGCCACCTTCTCTCTGCAAGAA 938
      |||
QY 257 ----GGATCGCTCTCAAGAGCGTCCGAGGAGGAGGACCAATCCTGTAGTGAAG 312
      |||
Db 939 ACTGGAGCTCGTCTCAAGAGCGTCCGAGGAGGAGGACCAATCCTGTAGTGAAG 998
      |||
QY 313 CTGGAGCCTCCGAGGCCATCCATCTCCTGACTTGGTACAGCCACTGCTACCCAT 372
      |||
Db 999 CTGGAGCCTCCGAGGCCATCCATCTCCTGACTTGGTACAGCCACTGCTACCCAT 1058
      |||
QY 373 TTCTGAGATGTTTCCCAAGTATCCACTGGGTCCCGCAGGCCCAAGTTTGGAGGCAGG 432
      |||
Db 1059 TTCTGAGATGTTTCCCAAGTATCCACTGGGTCCCGCAGGCCCAAGTTTGGAGGCAGG 1118
      |||
QY 433 GTTCCGCGAACAGCAGAGTCTCTGTGACCTGACAGGAGGCCGAGTTTGAACCCGGGGA 492
      |||
Db 1119 GTTCCGCGAACAGCAGAGTCTCTGTGACCTGACAGGAGGCCGAGTTTGAACCCGGGGA 1178
      |||
QY 493 GCAGAGCCAGGTGGCCACGCTACCAATGCGATTCATGTCACCGCGGCTCTATGACTAT 552
      |||
Db 1179 GCAGAGCCAGGTGGCCACGCTACCAATGCGATTCATGTCACCGCGGCTCTATGACTAT 1238
      |||
QY 553 CACTGGCAACATCTACATCTACAATGGACAGTACTGGGGGAGCACACCGGGTCTTGAGA 612
      |||
Db 1239 CACTGGCAACATCTACATCTACAATGGACAGTACTGGGGGAGCACACCGGGTCTTGAGA 1298
      |||
QY 613 CTCCCGAGTACCCCGAACCTCCATACCCCATTTCCC 649
      |||
Db 1299 CTCCCGAGTACCCCGAACCTCCATACCCCATTTCCC 1335
      |||

RESULT 4
AX409488 AX409488 2136 bp DNA linear PAT 14-JUN-2002
LOCUS Sequence 2135 from Patent WO0229103.
DEFINITION AX409488
ACCESSION AX409488
VERSION AX409488.1 GI:21442193
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2135 11-APR-2002;
GENE LOGIC INC (US)
FEATURES
source
Location/Qualifiers
i. .2136
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 446 a 706 c 608 g 376 t
ORIGIN
Query Match 81.8%; Score 531; DB 6; Length 2136;
Best Local Similarity 85.7%; Pred. No. 2.6e-129;
Matches 649; Conservative 0; Mismatches 0; Indels 108; Gaps 1;

QY 1 CGAGTGTACACTGCGAGCTACTTTCTGACTGCCCGCTGGCAGCTGAAGCCGAGTCAA 60
      |||
Db 579 CGAGTGTACACTGCGAGCTACTTTCTGACTGCCCGCTGGCAGCTGAAGCCGAGTCAA 638
      |||
QY 61 AGATGAAGTTGGGAAGGGTAACAACCACTGCGTCCCTCCAGGAGGAGGCACTTCCAGAA 120
      |||
Db 639 AGATGAAGTTGGGAAGGGTAACAACCACTGCGTCCCTCCAGGAGGAGGCACTTCCAGAA 698
      |||

```



```

/organism="Homo sapiens"
/db_xref="LocusID:4055"
/db_xref="taxon:9606"
/clone="MGC:22593 IMAGE:4703362"
/tissue_type="Lung"
/clone_lib="NIH_MGC_77"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
199, 1506
/codon_start=1
/product="lymphotoxin beta receptor (TNFR superfamily, member 3)"
/protein_id="AAH26262.1"
/db_xref="GI:20072213"
/translation="MLLPWATAPGLWPLVLGLLAAASQPAQVPPVASENOTCR
DQKEYEYRPHRCCSRCPGTVSAKSRIRDTVCATCAENSYNHNWHLSTCOLCR
PCDVMGLEIAPCTSKRTQCRQCFMCAWALECTHCELLSDCPTGETALKEDEV
GNNHVCPEKAGHFTSSPSARQCPHTRCENQGLVEAAPGTADQCTCKNPLEPLP
PENSMTMLMLAVLLPLAFLLATVFCISWKSRLKGLSLKKRRPQGEPPVAG
SWEPKRAHYFPDLVQPLLPISGDVSVSTGLPAFVLEAGVQQQSPDLDTREPOLE
PGEOSVAHGTNGIHVTGSMITIGNIYIYNGVPLGGPPGDLPATPEPPYPIPEG
DPGPPGLSTPHQDGRAHLAETEHGATPSNRGNRPNQFIHSD"
BASE COUNT 473 a 708 c 606 g 374 t
ORIGIN
Query Match 81.6%; Score 529.4; DB 9; Length 2161;
Best Local Similarity 85.6%; Pred. No. 7e-129;
Matches 648; Conservative 0; Mismatches 1; Indels 108; Gaps 1;
QY 1 CGAGTGACACACTGCGAGCTACTTTCTGACTGCGCCGCTGGCACTGAAGCGAGCTCAA 60
Db 609 CGAGTGACACACTGCGAGCTACTTTCTGACTGCGCCGCTGGCACTGAAGCGAGCTCAA 668
QY 61 AGATGAAGTTGGGAAGGTAAACACACTGCGTCCCTGCAAGCGAGCGCACTTCCAGAA 120
Db 669 AGATGAAGTTGGGAAGGTAAACACACTGCGTCCCTGCAAGCGCGGCACTTCCAGAA 728
QY 121 TACCTCTCCCCAGCCGCTGCCAGCCCCACACACAGGTGTGAGAACCAAGTCTGGT 180
Db 729 TACCTCTCCCCAGCCGCTGCCAGCCCCACACACAGGTGTGAGAACCAAGTCTGGT 788
QY 181 GGAGCGAGCTCCAGGCACTGCCAGTCCGACACACACCTGCAAAATCCATTAGAGCACT 240
Db 789 GGAGCGAGCTCCAGGCACTGCCAGTCCGACACACACCTGCAAAATCCATTAGAGCACT 848
QY 241 GCCCCCAGAGATGCA----- 256
Db 849 GCCCCCAGAGATGTCAGGAACCACTGTGCTGGCCGCTTCTGCTGCCACTGGCCCTCTT 908
QY 257 ----- 256
Db 909 TCTGCTCTTGGCCACGCTTCTCTCTGATCTGGAAGAGCCACCCCTTCTCTGCGAGAA 968
QY 257 ----GGATCGTGTCAAGAGGCTCGCAGGAGAGAGAGCCCAATCTGTAGTGAAG 312
Db 969 ACTGGGATCGTGTCAAGAGGCTCGCAGGAGAGAGAGCCCAATCTGTAGTGAAG 1028
QY 313 CTGGAGCCTCCGAAGCCCATCATACTTCCCTGACTTGTGACAGCACTGCTACCCAT 372
Db 1029 CTGGAGCCTCCGAAGCCCATCATACTTCCCTGACTTGTGACAGCACTGCTACCCAT 1088
QY 373 TTCTGGAGATGTTTCCCGAGTATCCACTGGGCTCCCGGAGCCGCCAGTTTGGAGGAGG 432
Db 1089 TTCTGGAGATGTTTCCCGAGTATCCACTGGGCTCCCGGAGCCGCCAGTTTGGAGGAGG 1148
QY 433 GGTCCCGACAGCAGAGTCTCTGAGACTGACACAGGAGCGCGAGTTGGAACCCGGGGA 492
Db 1149 GGTCCCGACAGCAGAGTCTCTGAGACTGACACAGGAGCGCGAGTTGGAACCCGGGGA 1208
QY 493 GCAGAGCAGGTGGCCACAGTACCAATGCAATTCATGTCACCGCGGGTCTATGACTAT 552
Db 1209 GCAGAGCAGGTGGCCACAGTACCAATGCAATTCATGTCACCGCGGGTCTATGACTAT 1268

```

```

QY 553 CACTGGCAACATCTACATCTACATGACACAGTACTGGGGGACACCCGGTCTCTGGAGA 612
Db 1269 CACTGGCAACATCTACATCTACATGACACAGTACTGGGGGACACCCGGTCTCTGGAGA 1328
QY 613 CACTCCAGCTACCCCGAAGCTCCATACCCCATTCOC 649
Db 1329 CACTCCAGCTACCCCGAAGCTCCATACCCCATTCOC 1365

RESULT 8
MUSLYMPHOB 1614 bp mRNA linear ROD 24-MAY-1996
LOCUS Mus musculus lymphotoxin-beta receptor gene, complete cds.
DEFINITION L38423
VERSION L38423.1 GI:600222
KEYWORDS lymphotoxin-beta receptor; transmembrane protein.
SOURCE Mus musculus.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1614)
AUTHORS Nakamura,T., Tashiro,K., Nazarea,M., Nakano,T., Sasayama,S. and
Honjo,T.
TITLE The murine lymphotoxin-beta receptor cDNA: isolation by the signal
sequence trap and chromosomal mapping
JOURNAL Genomics 30 (2), 312-319 (1995)
MEDLINE 96163885
PubMed 8586432
FEATURES
source
1..1614
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_line="ST-2"
186..1433
/note="putative"
/codon_start=1
/product="lymphotoxin-beta receptor"
/protein_id="AAB00846.1"
/db_xref="GI:600223"
/translation="MRLPRASSPCGLWGLPLLLGLLVAQPLVPPYRIENOTCM
DQKEYEYRPHDVCSCRPQPFVAVCSRODTCKTKPHNSYNHNWHLSTCOLCR
PCDVLGFEEVAFCTSDRAECRCQPMSCVYLDNECVHCEERLVLCOQPEAEVTD
EIMDTCNVCPKPHQNTSSPRARQCPHTRCENQGLVEAAPGTADQCTCKNPPPEP
GAMLLAILLLPLFLFTTFLACAMRHPSLCRLKGLTLKKRHPGEESSPPCPAPRAD
PHFPLAEEPLPMSGLSPSPAPPTAPSLVEEVQQQSLVQARLEAEPEHGGVA
HGANGIHVTGSGVTGNIYIYNGVPLGTRGPGDPPADPEPPYPTPERGAPGPSLS
"TPYQEDGKAHLAETETLSCQDI."
BASE COUNT 316 a 537 c 455 g 306 t
ORIGIN
Query Match 41.6%; Score 270; DB 10; Length 1614;
Best Local Similarity 65.8%; Pred. No. 1.9e-60;
Matches 467; Conservative 0; Mismatches 165; Indels 78; Gaps 2;
QY 17 GAGTACTTTCTGACTGCCCGCTGGCACTGAAGCGAGCTCAAAGTGAAGTTGGGAAG 76
Db 618 GAGCGGCTTCTACTCTGCCAGCTGGCACAGAAGCCGAGCTCAGATGAATATGAT 677
QY 77 GGTAAACAACACTGCGTCCCTCAAGCGAGGCACTTCCAGATACTCTCCCCAGC 136
Db 678 ACTGAGCTCAACTGTGTCCCTGTAGCCGGGACACTTCCAGAACACTTCTCCCTCTCGA 737
QY 137 GCCCGCTGCCAGCCCCACACAGGTGTGAGAACCAAGGTCTGTGGAGGAGCTCCAGGC 196
Db 738 GCCCGCTGTCAACCCCATACAGATGTGAGATCCAGGCGCTGTGGAGGAGCTCCAGGT 797
QY 197 ACTGCCAGTCCGACACACACTGCAAAATTCATTAGAGCACTGCCCCCAGAGATGCA 256
Db 798 ACCTCCTACTCGGATACCATCTGTAAAAATCCCCAGAGCGAGCAATGCTACTGCT 857
QY 257 GGATCGTCTCTCAAG-----AGGCGTCCGAGGG 285
Db 858 GCCATCCTGCTGCTGCTGGTCTCTTTCTGCTCTTACCACACTGCTGTCGCTGTGCTCTG 917

```


TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					
26					
27					
28					
29					
30					
31					
32					
33					
34					
35					
36					
37					
38					
39					
40					
41					
42					
43					
44					
45					
46					
47					
48					
49					
50					
51					
52					
53					
54					
55					
56					
57					
58					
59					
60					
61					
62					
63					
64					
65					
66					
67					
68					
69					
70					
71					
72					
73					
74					
75					
76					
77					
78					
79					
80					
81					
82					
83					
84					
85					
86					
87					
88					
89					
90					
91					
92					
93					
94					
95					
96					
97					
98					
99					
100					

Direct Submission
Submitted (22-OCT-1998) Department of Molecular Biology, Albert
Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY
10461, USA
3 (bases 1 to 140026)
Montgomery, K.T., Lau, S.T. and Kucherlapati, R.
Direct Submission
Submitted (20-OCT-2000) Department of Molecular Genetics, Albert
Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY
10461, USA
On Oct 20, 2000 this sequence version replaced gi:3779004.
-----Genome Center:
Albert Einstein College of Medicine
Code: AECOM
Web site: <http://sequence.aecom.yu.edu/chrl2/>
Contact: ihane@sequence.aecom.yu.edu

CLONE LENGTH: This sequence represents the entire insert of this clone unless otherwise noted. If there are overlapping clones, the overlaps are noted in the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSS are identified using ePCR (Genome Res. 7:541-550). Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human sequences.

Genes and Regions of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to EST and cDNA sequences in Unigene. Genes demonstrate at least two exons flanked by consensus splice sites that maintain sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Attempts are made to complete double stranded sequence for all regions. All sequence is completed to a standard of coverage with a minimum of 3 reads with no ambiguities. If the sequence coverage for a region does not meet this standard, it is indicated in the annotation as Low Coverage. Low coverage linkages are verified by PCR product size verification or verification of forward and reverse reads from clones which span the low coverage area.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated average error rate is less than 1 per 10,000 bases using the Consed quality parameters. Regions that do not meet this requirement are annotated as Low Quality.

-----Summary Statistics

Center project name: RPI-102E24
Sequencing vector: M13
Chemistry: Dye-terminator Big Dye; 100%
Assembly program: Phrap version 0.990319
Contig length: 14026
Fraction of Phrap value < 40: 0.0426
Error Rate in Consed: 0.56 per 10,000 bp
Number of N's in consensus: 1

----- Distribution of quality < 40 Bases:

[illegible]

FEATURES	Location/Qualifiers
source	1..140026
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="12"
	/clone="RP1-102E24"
repeat_region	1..51
	/rpt_family="AluJb"
repeat_region	52..76
	/rpt_family="(CAAA)n"
repeat_region	complement(254..367)
	/rpt_family="L1MB8"
repeat_region	complement(377..649)
	/rpt_family="L1MA8"
repeat_region	complement(650..941)
	/rpt_family="AluSc"
repeat_region	complement(983..1090)
	/rpt_family="L1MB8"
repeat_region	complement(1110..1339)
	/rpt_family="AluSx"
repeat_region	complement(1341..1354)
	/rpt_family="L1MB8"
repeat_region	complement(1355..1652)
	/rpt_family="AluSx"
repeat_region	complement(1653..1787)
	/rpt_family="L1MB8"
repeat_region	complement(1797..2090)
	/rpt_family="AluJo"
repeat_region	complement(2241..2533)
	/rpt_family="AluSg"
repeat_region	3376..3664
	/rpt_family="AluSg"
repeat_region	3825..3850
	/rpt_family="AT-rich"
gene	complement(join(4534..485517..5666,5694..5849,56889..7029))
	/gene="Plakophilin 2 (PKP7134, 7432
repeat_region	/rpt_family="AluSc"
repeat_region	complement(7442..7588)
	/rpt_family="L1M4"
repeat_region	8060..8245
	/rpt_family="L1MB2"
repeat_region	8246..8360
	/rpt_family="HAL1"
repeat_region	8399..8563
	/rpt_family="AluSg"
repeat_region	8564..9933
	/rpt_family="SVA"
repeat_region	9934..10088
	/rpt_family="L1MD2"
repeat_region	10107..10188
	/rpt_family="FLAM_A"
repeat_region	10194..10975
	/rpt_family="L1MD2"
repeat_region	10976..11283
	/rpt_family="AluY"
repeat_region	complement(11303..11589)
	/rpt_family="AluJb"
repeat_region	complement(11678..11692)
	/rpt_family="FRAM"
repeat_region	complement(11693..11995)
	/rpt_family="AluSx"
repeat_region	complement(11996..12157)
	/rpt_family="FRAM"
repeat_region	12547..12899
	/rpt_family="MER104"
repeat_region	complement(12900..13198)
	/rpt_family="AluSx"
repeat_region	13199..13254
	/rpt_family="MER104"

```

repeat_region complement(13309..13472)
STS /rpt_family="L2"
13793..14017
/standard_name="GDB:384731"
repeat_region complement(16498..17181)
/rpt_family="THEIA-int"
repeat_region complement(17193..17563)
/rpt_family="THEIA-int"
repeat_region complement(17564..17915)
/rpt_family="THEIA"
repeat_region complement(18210..18342)
/rpt_family="FLAM-A"
gene join(18391..18519,18744..18902,19803..19853)
/feature="Vesicle-associated membrane protein 1 (SVB1)"
19248..19275
repeat_region 19276..19351
/rpt_family="CA)n"
repeat_region 19352..19475
/rpt_family="LIMed"
repeat_region complement(19296..19475)
/rpt_family="LIMCa"
repeat_region 19914..19933
/rpt_family="(CCTG)n"
STS 20287..20651
/standard_name="GDB:214816"
STS 20290..20527
/standard_name="sts-M36200"
STS 20459..20650
/standard_name="sts-M36200"
STS 20841..21018
/standard_name="SHGC-557"
STS 20971..21045
/standard_name="D12S111E"
STS 22216..22330
/standard_name="Cda1eb01"
STS 22518..22670
/standard_name="stS4862"
gene complement(join(22595..22710,23806..23889,25797..26099,
27000..27338,31028..31297,31447..31677,32432..32495))
Query Match 36.3%; Score 235.4; DB 9; Length 140026;
Best Local Similarity 99.6%; Pred. No. 2.3e-51;
Matches 236; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 278 CGCAGGAGGAGGACCAATCTCTAGCTGGAGCTGGAGCCCTCCGAAGCCCATCCA 337
|||||
Db 94638 CTGACGGAGAGGAGGACCAATCTCTAGCTGGAGCTGGAGCCCTCCGAAGCCCATCCA 94579

QY 338 TACTTCCCTGACTGGTACAGCCACTGCTACCCATTTCTGGAGATGTTTCCCCAGTATCC 397
|||||
Db 94578 TACTTCCCTGACTGGTACAGCCACTGCTACCCATTTCTGGAGATGTTTCCCCAGTATCC 94519

QY 398 ACTGGCTCCCGCAGCCCGAGTTTGGAGGAGGAGGTCGCGCAACAGCAGAGTCTCTG 457
|||||
Db 94518 ACTGGCTCCCGCAGCCCGAGTTTGGAGGAGGAGGTCGCGCAACAGCAGAGTCTCTG 94459

QY 458 GACCTGACAGGAGCCCGAGTTGGAACCGGGGAGCAGAGCAGGTGCCACCGGT 514
|||||
Db 94458 GACCTGACAGGAGCCCGAGTTGGAACCGGGGAGCAGAGCAGGTGCCACCGGT 94402

RESULT 11
LOCUS HUMTUMNEC 1605 bp mRNA linear PRI 14-JAN-1995
DEFINITION Homo sapiens (clone NCPI8) tumor necrosis factor receptor related
protein mRNA, complete exon and repeat region.
ACCESSION L04489
VERSION L04489.1 GI:340022
KEYWORDS tumor necrosis factor receptor related protein.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1605)

```

```

AUTHORS Baens,M., Chaffanet,M., Cassiman,J.J., van den Berghe,H. and
Marynen,P.
TITLE Construction and evaluation of a hncDNA library of human 12p
transcribed sequences derived from a somatic cell hybrid
JOURNAL Genomics 16 (1), 214-218 (1993)
MEDLINE 93252381
PUBMED 8486360
FEATURES
source Location/Qualifiers
1..1605
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="12p13"
/cell_line="M28 somatic cell hybrid"
/tissue_lib="hncdna tc651"
1..1605
/gene="TNFR"
repeat_region 1..105
/partial
/feature="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
250..444
/gene="TNFR"
exon /feature="L04270 sequence homologue"
/feature="experimental"
BASE COUNT 472 a 394 c 323 g 416 t
ORIGIN

Query Match 30.5%; Score 198; DB 9; Length 1605;
Best Local Similarity 100.0%; Pred. No. 1.8e-41;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 AGATCAAGTTGGAGGTTAACAACCACTGCGTCCCTGCAAGCAGGCGCTCCAGAA 120
|||||
Db 248 AGATCAAGTTGGAGGTTAACAACCACTGCGTCCCTGCAAGCAGGCGCTCCAGAA 307

QY 121 TACCTCTCTCCCGCAGCGCGCTGCCAGCCGCGCAGGCTGTGAGAACCAAGGTCTGT 180
|||||
Db 308 TACCTCTCTCCCGCAGCGCGCTGCCAGCCGCGCAGGCTGTGAGAACCAAGGTCTGT 367

QY 181 GGAGCAGCTCCAGCAGTCCCGCAGTCCGAGCACACACCTGCAAAATCCATTAGAGCCCT 240
|||||
Db 368 GGAGCAGCTCCAGCAGTCCCGCAGTCCGAGCACACACCTGCAAAATCCATTAGAGCCCT 427

QY 241 GCCCCAGAGATGTCCAGG 258
|||||
Db 428 GCCCCAGAGATGTCCAGG 445

RESULT 12
LOCUS AC128082/c 178228 bp DNA linear HTG 19-JUL-2002
DEFINITION Rattus norvegicus clone CH230-362C16, *** SEQUENCING IN PROGRESS
ACCESSION AC128082
VERSION AC128082.1 GI:21908679
KEYWORDS HTG; HTGS; PHASEL.
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 178228)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayale,M., Banks,F.,
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowle,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Caron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleaveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,

```

Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
 Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,
 Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
 Li, J., Li, E., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louseghed, H.,
 Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
 Massey, E., Mahoney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G.,
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
 Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G.,
 Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I.,
 Sodergren, E., Sonaik, I., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wleczek, R., Woodson, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstein, G., and Gibbs, R.
 Direct Submission
 2 (bases 1 to 178228)
 Unpublished
 Worley K.C.
 Direct Submission
 Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

 Center: Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu

 Project Information
 Center project name: GZQW
 Center clone name: CH230-362C16

 Summary Statistics
 Sequencing vector: Plasmid
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 116771 bases at least Q40
 Consensus quality: 123909 bases at least Q30
 Consensus quality: 129462 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 64 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1173: contig of 1173 bp in length
 * 1174: 1273: gap of unknown length
 * 1274: 2435: contig of 1162 bp in length
 * 2436: 2535: gap of unknown length
 * 2536: 3567: contig of 1032 bp in length
 * 3568: 3667: gap of unknown length
 * 3668: 5230: contig of 1563 bp in length
 * 5231: 5330: gap of unknown length
 * 5331: 6732: contig of 1402 bp in length

6733: 6832: gap of unknown length
 6833: 7990: contig of 1158 bp in length
 7991: 8090: gap of unknown length
 8091: 9641: contig of 1551 bp in length
 9642: 9741: gap of unknown length
 9742: 10776: contig of 1035 bp in length
 10777: 10876: gap of unknown length
 10877: 11905: contig of 1029 bp in length
 11906: 12005: gap of unknown length
 12006: 13202: contig of 1315 bp in length
 13207: 13420: gap of unknown length
 13421: 14778: contig of 1258 bp in length
 14779: 14778: gap of unknown length
 14780: 15851: contig of 1073 bp in length
 15852: 15951: gap of unknown length
 15952: 17593: contig of 1642 bp in length
 17594: 17993: gap of unknown length
 17994: 18911: contig of 1218 bp in length
 18912: 19011: gap of unknown length
 19013: 20145: contig of 1134 bp in length
 20146: 20245: gap of unknown length
 20246: 21556: contig of 1311 bp in length
 21557: 21556: gap of unknown length
 21558: 22715: contig of 1059 bp in length
 22716: 22815: gap of unknown length
 22817: 23949: contig of 1134 bp in length
 23950: 24049: gap of unknown length
 24050: 25331: contig of 1282 bp in length
 25332: 25432: gap of unknown length
 25433: 26617: contig of 1186 bp in length
 26618: 26717: gap of unknown length
 26719: 28435: contig of 1718 bp in length
 28436: 28535: gap of unknown length
 28536: 30660: contig of 2125 bp in length
 30661: 30760: gap of unknown length
 30761: 32263: contig of 1503 bp in length
 32264: 32363: gap of unknown length
 32365: 34165: contig of 1802 bp in length
 34166: 34265: gap of unknown length
 34266: 35302: contig of 1037 bp in length
 35303: 35402: gap of unknown length
 35403: 37282: contig of 1880 bp in length
 37283: 37382: gap of unknown length
 37384: 39160: contig of 1778 bp in length
 39161: 39260: gap of unknown length
 39261: 40880: contig of 1620 bp in length
 40881: 40980: gap of unknown length
 40982: 42436: contig of 1456 bp in length
 42437: 42536: gap of unknown length
 42537: 44381: contig of 1845 bp in length
 44382: 44481: gap of unknown length
 44483: 46391: contig of 1910 bp in length
 46392: 46491: gap of unknown length
 46492: 48409: contig of 1918 bp in length
 48410: 48509: gap of unknown length
 48510: 49931: contig of 1422 bp in length
 49932: 50032: gap of unknown length
 50033: 51371: contig of 1340 bp in length
 51372: 51471: gap of unknown length
 51472: 53362: contig of 1891 bp in length
 53363: 53462: gap of unknown length
 53464: 55270: contig of 1708 bp in length
 55271: 55797: gap of unknown length
 55798: 56797: contig of 1427 bp in length
 56799: 56998: gap of unknown length
 56999: 58037: contig of 1240 bp in length
 58038: 58137: gap of unknown length
 58138: 60361: contig of 2224 bp in length
 60362: 62335: gap of unknown length
 62336: 62435: gap of unknown length
 62436: 64906: contig of 2471 bp in length
 64907: 65006: gap of unknown length

```

* 65007 67475: contig of 2469 bp in length
* 67476 67575: gap of unknown length
* 67576 69653: contig of 2078 bp in length
* 69654 69753: gap of unknown length
* 69754 73830: contig of 4077 bp in length
* 73831 73930: gap of unknown length
* 73931 77047: contig of 3117 bp in length
* 77048 77147: gap of unknown length
* 77148 80438: contig of 3291 bp in length
* 80439 80538: gap of unknown length
* 80539 83532: contig of 2994 bp in length
* 83533 83632: gap of unknown length
* 83633 85754: contig of 2122 bp in length
* 85755 85854: gap of unknown length
* 85855 89203: contig of 3349 bp in length
* 89204 89303: gap of unknown length
* 89304 92865: contig of 3562 bp in length
* 92866 92965: gap of unknown length
* 92966 95763: contig of 2798 bp in length
* 95764 95863: gap of unknown length
* 95864 99798: contig of 3835 bp in length
* 99799 104352: contig of 4554 bp in length
* 104353 104452: gap of unknown length
* 104453 108165: contig of 3713 bp in length
* 108166 108265: gap of unknown length
* 108266 111945: contig of 3680 bp in length
* 111946 112045: gap of unknown length
* 112046 116301: contig of 4256 bp in length
* 116302 116401: gap of unknown length
* 116402 120825: contig of 4424 bp in length
* 120826 120925: gap of unknown length

Query Match 18.1%; Score 117.6; DB 2; Length 178228;
Best Local Similarity 72.2%; Pred. No. 2.7e-20;
Matches 153; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 313 CTGGAGCCTCGAAGGCCCATCTACATCTCCCTGACTTGATGACGACCTGCTACCCAT 372
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 152659 CTGCCAGCCTCAAGAGCCCAACACATTTCCGTGACCTGGCAGACCACTCTACCCAT 152600

QY 373 TTCTGAGATGTTCCCAAGTATCCACTGGCTGCCGAGCCGACCCAGTTTGGAGCAGG 432
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 152599 GTCTGGAGACTTGTCCCAAGCCCTCTGTGAGCCGCCCTCTCTGGAGGAGT 152540

QY 433 GTGCGCGCAACAGCAGAGCTCTGACCTGACGAGGAGCGCAGTGGACCCGGGCA 492
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 152539 GTGCTACAGCAGCATAGTCCCTGATCCAGGCGGAGCTGGAGCTGAGCCTGGGA 152480

QY 493 GCAGAGCCAGGTGGCCCGACGCTACCAATGGCA 524
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 152479 ACATGCCAGGTGGCCCGACGCTGAGTGGGCA 152448

RESULT 13
HS27B9F
LOCUS HS27B9F 323 bp DNA linear PRI 19-OCT-1995
DEFINITION H.sapiens CpG island DNA genomic Msel fragment, clone 27B9, reverse
read cp927b9.ftla.
ACCESSION Z60529
VERSION Z60529.1 GI:1032633
KEYWORDS CpG island; genomic Msel fragment.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
MacDonald,M., Huckle,E., Wilkinson,P. and Micklem,G.
Direct Submission
Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
CB10 1RQ, England. E-mail contact: humquery@sanger.ac.uk
REFERENCE 2 (bases 1 to 323)
AUTHORS Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.

Purification of CpG islands using a methylated DNA binding column
Nat. Genet. 6 (3), 236-244 (1994)
94282070
8012384
Vector: pGEM-5zf(-)
Clones are available from the UK MRC Human Genome Mapping Project
Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
http://www.hgmp.mrc.ac.uk/ for details
or contact: biohelp@hgmp.mrc.ac.uk.

FEATURES
source
1..323
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="27b9"
/sex="male"
/tissue_type="blood"
/clone_lib="CGI-1"
/dev_stage="adult"
BASE COUNT 64 a 123 c 75 g 56 t 5 others
ORIGIN

Query Match 15.3%; Score 99; DB 9; Length 323;
Best Local Similarity 92.7%; Pred. No. 2.5e-15;
Matches 102; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 61 AGATGAAGTTGGAGGGTAACAACACCTGCGTCCCTGCAAGGCGGCACTTCCAGAA 120
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 75 AGATGAAGTTGGAGGGTAACAACACCTGCGTCCCTGCAAGGCGGCACTTCCAGAA 134

QY 121 TACCTCTCCCCAGCGCCGCTGCCAGCCCAACACAGGTGTGAGAACC 170
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 135 TACCTCTCCCCAGCGCCGCTGCCAGCCCAACACAGGTGTGAGTGCAGC 184

RESULT 14
HS27B9F/c
LOCUS HS27B9F/c 193 bp DNA linear PRI 19-OCT-1995
DEFINITION H.sapiens CpG island DNA genomic Msel fragment, clone 27B9, forward
read cp927b9.ftla.
ACCESSION Z60528
VERSION Z60528.1 GI:1032632
KEYWORDS CpG island; genomic Msel fragment.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
MacDonald,M., Huckle,E., Wilkinson,P. and Micklem,G.
Direct Submission
Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
CB10 1RQ, England. E-mail contact: humquery@sanger.ac.uk
REFERENCE 2 (bases 1 to 193)
AUTHORS Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
Purification of CpG islands using a methylated DNA binding column
Nat. Genet. 6 (3), 236-244 (1994)
94282070
8012384
Vector: pGEM-5zf(-)
Clones are available from the UK MRC Human Genome Mapping Project
Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
http://www.hgmp.mrc.ac.uk/ for details
or contact: biohelp@hgmp.mrc.ac.uk.

FEATURES
source
1..193
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="27b9"
/sex="male"
/tissue_type="blood"
/clone_lib="CGI-1"
/dev_stage="adult"
BASE COUNT 36 a 60 c 48 g 44 t 5 others
ORIGIN

```

```

TITLE Purification of CpG islands using a methylated DNA binding column
JOURNAL Nat. Genet. 6 (3), 236-244 (1994)
MEDLINE 94282070
PUBMED 8012384
COMMENT Vector: pGEM-5zf(-)
Clones are available from the UK MRC Human Genome Mapping Project
Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
http://www.hgmp.mrc.ac.uk/ for details
or contact: biohelp@hgmp.mrc.ac.uk.

```

FEATURES

source

1..323
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="27b9"
 /sex="male"
 /tissue_type="blood"
 /clone_lib="CGI-1"
 /dev_stage="adult"

BASE COUNT 64 a 123 c 75 g 56 t 5 others
 ORIGIN

Query Match 15.3%; Score 99; DB 9; Length 323;

Best Local Similarity 92.7%; Pred. No. 2.5e-15;
 Matches 102; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 61 AGATGAAGTTGGAGGGTAACAACACCTGCGTCCCTGCAAGGCGGCACTTCCAGAA 120

Db 75 AGATGAAGTTGGAGGGTAACAACACCTGCGTCCCTGCAAGGCGGCACTTCCAGAA 134

QY 121 TACCTCTCCCCAGCGCCGCTGCCAGCCCAACACAGGTGTGAGAACC 170

Db 135 TACCTCTCCCCAGCGCCGCTGCCAGCCCAACACAGGTGTGAGTGCAGC 184

RESULT 14

HS27B9F/c

LOCUS HS27B9F/c 193 bp DNA linear PRI 19-OCT-1995

DEFINITION H.sapiens CpG island DNA genomic Msel fragment, clone 27B9, forward

read cp927b9.ftla.

ACCESSION Z60528

VERSION Z60528.1 GI:1032632

KEYWORDS CpG island; genomic Msel fragment.

SOURCE Homo sapiens.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

MacDonald,M., Huckle,E., Wilkinson,P. and Micklem,G.

Direct Submission

Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,

CB10 1RQ, England. E-mail contact: humquery@sanger.ac.uk

REFERENCE 2 (bases 1 to 193)

AUTHORS Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.

Purification of CpG islands using a methylated DNA binding column

Nat. Genet. 6 (3), 236-244 (1994)

94282070

8012384

Vector: pGEM-5zf(-)

Clones are available from the UK MRC Human Genome Mapping Project

Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:

http://www.hgmp.mrc.ac.uk/ for details

or contact: biohelp@hgmp.mrc.ac.uk.

FEATURES

source

1..193

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="27b9"

/sex="male"

/tissue_type="blood"

/clone_lib="CGI-1"

/dev_stage="adult"

BASE COUNT 36 a 60 c 48 g 44 t 5 others

ORIGIN

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:05:15 ; Search time 100.083 Seconds
(without alignments)
12938.300 Million cell updates/sec

Title: US-09-917-372-2_COPY_804_1378

Perfect score: 575

Sequence: 1 ctgtgtgagcagctccagg.....cactgtgtggtccacaccctc 575

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

- 1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	494	85.9	1594	ABQ55003	Human ovarian anti
2	494	85.9	2136	ABN95637	Gene #2135 used to
3	494	85.9	2136	ABK64502	Human benign prost
4	494	85.9	2136	ABL64078	Breast cancer rela
5	494	85.9	2136	ABL64384	Stomach cancer rel
6	253.8	44.1	281	ABL82623	Human ovarian canc
7	251.2	43.7	289	ABL82688	Human ovarian canc
8	61.2	10.6	481	ABL81903	Human ovarian canc
9	36.4	6.3	333	AAV7224	Human anti-GPIIb/I

c 10	36.4	6.3	788	22	AAH07782	Human cDNA clone (
c 11	2497	6.3	2497	22	AAI93910	Human stomach canc
c 12	36.4	6.3	2497	22	AAH18071	Human cDNA sequenc
c 13	36.4	6.3	3290	21	AAC59813	Human secreted pro
c 14	36.4	6.3	3298	22	AAS40920	cDNA encoding nove
c 15	36.2	6.3	450	23	ABV38197	Human prostate exp
c 16	36	6.3	3565	22	AAH40928	cDNA encoding nove
c 17	35.6	6.2	4784	22	AAH98417	Murine EST-derived
c 18	35.6	6.2	4791	22	AAS22677	Human cDNA encodin
c 19	35.4	6.2	10336	22	ABAL7315	Human nervous syst
c 20	35.4	6.2	20813	22	ABAL7314	Human nervous syst
c 21	35.2	6.1	4428	22	AAAD06574	Bovine alpha(III)
c 22	35.2	6.1	4428	22	AAAD06575	Bovine alpha(III)
c 23	35	6.1	1576	21	AAAC77136	Human ORF ORF2691
c 24	34.6	6.0	1521	23	AA571133	DNA encoding novel
c 25	34.6	6.0	2005	24	ABL95574	Human angiogenesis
c 26	34.6	6.0	2005	24	ABL88085	Human PRO7223 cDNA
c 27	34.6	6.0	2024	21	AAA07697	Human collectin en
c 28	34.6	6.0	2181	22	AAI60628	Human polynucleoti
c 29	34.6	6.0	2262	24	ABA97932	Human scavenger re
c 30	34.6	6.0	2318	22	AAI58842	Human polynucleoti
c 31	34.6	6.0	2628	22	AAH43036	Nucleotide sequenc
c 32	34.6	6.0	2641	22	AAC68903	Human EXMAD-14 cod
c 33	34.6	6.0	2929	20	AAQ92072	Human polynucleoti
c 34	34.6	6.0	2930	20	AAV55746	Human secreted pro
c 35	34.6	6.0	3535	22	AAH47802	Chimeric CBD-fused
c 36	34.6	6.0	5692	22	AAF32248	Streptomyces sp. c
c 37	34.4	6.0	2768	24	ABN89352	Mouse E-cadherin e
c 38	34.4	6.0	2768	24	ABN89352	Mouse E-cadherin e
c 39	34.2	5.9	268	13	AAQ25135	DR1. Synthetic.
c 40	34.2	5.9	874	21	AAC62307	Contig HEPATOX7, a
c 41	34	5.9	3474	14	AAQ39018	Sequence of pure m
c 42	34	5.9	3474	18	AAT74012	Maize optimised-B.
c 43	34	5.9	3474	19	AAV16191	Maize optimised DN
c 44	34	5.9	11244	22	AAK74606	Human immune/haema
c 45	34	5.9	15296	22	AAK74605	Human immune/haema

ALIGNMENTS

RESULT 1

ABQ55003	ID	ABQ55003 standard; cDNA; 1594 BP.
XX	AC	ABQ55003;
XX	DT	22-AUG-2002 (first entry)
XX	DE	Human ovarian antigen HSABJ44 cDNA, SEQ ID NO:883.
XX	DE	Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW	KW	ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW	KW	infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW	KW	PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW	KW	inflammatory condition; immune disorder; blood disorder;
KW	KW	cardiovascular disorder; respiratory disorder; neurological disorder;
KW	KW	gastrointestinal disorder; urinary system disorder; drug screening;
KW	KW	gene therapy; chromosome mapping; forensic analysis;
KW	KW	antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW	KW	antiinflammatory; gynaecological; reproductive; chromosome 12p13;
KW	KW	gene; ss.
XX	OS	Homo sapiens.
XX	PN	WO200200677-A1.
XX	PD	03-JAN-2002.
XX	PF	07-JUN-2001; 2001WO-US18569.
XX	PR	07-JUN-2000; 2000US-209467P.

QY 82 GGATCGCTGCTCAAGAGGCGTCCGAGGAGAGGAGGCCCAATCTGTAGCTGGAAGCTGG 141
|||||
Db 943 GGATCGCTGCTCAAGAGGCGTCCGAGGAGAGGAGGCCCAATCTGTAGCTGGAAGCTGG 1002
|||||
QY 142 GAGCCTCCGAGGCCCATCATCTATCTCCCTGACTGTGTACAGCCACTGCTACCCATTCT 201
|||||
Db 1003 GAGCCTCCGAGGCCCATCATCTATCTCCCTGACTGTGTACAGCCACTGCTACCCATTCT 1062
|||||
QY 202 GGAGATGTTTCCCGACATCATCTGAGTCCCGCCGAGCCCGCCAGTTTGGAGGAGGGTG 261
|||||
Db 1063 GGAGATGTTTCCCGACATCATCTGAGTCCCGCCGAGCCCGCCAGTTTGGAGGAGGGTG 1122
|||||
QY 262 CCACAACAGCAGAGTCTCTGAGCTGACACAGGAGCGCGAGTTGGAACCGGGGAGCAG 321
|||||
Db 1123 CCACAACAGCAGAGTCTCTGAGCTGACACAGGAGCGCGAGTTGGAACCGGGGAGCAG 1182
|||||
QY 322 AGCAGGTGGCCGACGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 381
|||||
Db 1183 AGCAGGTGGCCGACGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1242
|||||
QY 382 GGCAACATCTACATCTACATCTACATCTACATCTACATCTACATCTACATCTACATCT 441
|||||
Db 1243 GGCAACATCTACATCTACATCTACATCTACATCTACATCTACATCTACATCTACATCT 1302
|||||
QY 442 CCAGCTACCCCGAACCCTCCATACCCCATTCGCGAAGAGGGGACCCCTGGCCCTCCCGGG 501
|||||
Db 1303 CCAGCTACCCCGAACCCTCCATACCCCATTCGCGAAGAGGGGACCCCTGGCCCTCCCGGG 1362
|||||
QY 502 CTCTCTACACCCACACAGGAGATGGCAAGCTTGGCACTGATGATGATGATGATGATGATGAT 561
|||||
Db 1363 CTCTCTACACCCACACAGGAGATGGCAAGCTTGGCACTGATGATGATGATGATGATGATGAT 1422
|||||
QY 562 GGTGCCACACCCCTC 575
|||||
Db 1423 GGTGCCACACCCCTC 1436
|||||

RESULT 3

ABK64502

ID ABK64502 standard; DNA; 2136 BP.

XX
AC ABK64502;

XX 18-JUN-2002 (first entry)

XX Human benign prostatic hyperplasia gene #397.

DE Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.
KW Homo sapiens.XX
OS Wo200212440-A2.XX
PN 14-FEB-2002.XX
PD 07-AUG-2001; 2001WO-US24708.XX
PF 07-AUG-2000; 2000US-223323P.XX
PR 05-JUN-2001; 2001US-0873319.XX
PA (GENE-) GENE LOGIC INC.XX
PA (NISB) JAPAN TOBACCO INC.XX
PI Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;
XX WPI; 2002-257476/30.XX
DR Identifying drugs for and diagnosing benign prostatic hyperplasia, by
XX detecting expression levels of one or more genes in prostate cells from
XX patient that are differentially regulated compared to normal prostate
XX cells

PS Disclosure; Page 229-230; 444pp; English.

XX The invention relates to a method of diagnosing (I) the onset or
XX progression of benign prostatic hyperplasia (BPH), or screening (II) for
XX identifying an agent that modulates the onset or progression of BPH.
XX The method is based on changes in gene expression in BPH tissue isolated
XX from patients exhibiting different clinical states of prostate
XX hyperplasia as compared to normal prostate tissue. (I) comprises
XX detecting the expression levels of one or more genes in prostate cells
XX from the subject that are differentially regulated compared to normal
XX prostate cells. (II) comprises preparing a first gene expression profile
XX of BPH cells or BPH-like cell population, exposing the cells to the
XX agent, preparing a second gene expression profile of the agent exposed
XX cells, and comparing the first and second gene expression profiles.
XX (I) is useful for diagnosing the onset or progression of BPH. (II) is
XX useful for identifying an agent that modulates the onset or progression
XX of BPH. The methods are useful to present information identifying
XX the expression level in a tissue or cells, by comparing the expression
XX level of genes given in the specification in the tissue or cells to the
XX level of expression of gene in the database, and displaying the
XX expression levels of at least one gene in the tissue or cell sample
XX compared to the expression level in BPH. Agents using (II) are useful for
XX treating BPH or prostate cancer. ABK64106-ABK64860 represent human
XX benign prostatic hyperplasia gene sequences of the invention.

SQ Sequence 2136 BP; 446 A; 706 C; 608 G; 376 T; 0 other;

Query Match 85.9%; Score 494; DB 24; Length 2136;

Best Local Similarity 100.0%; Pred. No. 1.6e-127;

Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 GGATCGCTGCTCAAGAGGCGTCCGAGGAGAGGAGGCCCAATCTGTAGCTGGAAGCTGG 141
|||||Db 943 GGATCGCTGCTCAAGAGGCGTCCGAGGAGAGGAGGCCCAATCTGTAGCTGGAAGCTGG 1002
|||||QY 142 GAGCCTCCGAGGCCCATCATCTATCTCCCTGACTGTGTACAGCCACTGCTACCCATTCT 201
|||||Db 1003 GAGCCTCCGAGGCCCATCATCTATCTCCCTGACTGTGTACAGCCACTGCTACCCATTCT 1062
|||||QY 202 GGAGATGTTTCCCGACATCATCTGAGTCCCGCCGAGCCCGCCAGTTTGGAGGAGGGTG 261
|||||Db 1063 GGAGATGTTTCCCGACATCATCTGAGTCCCGCCGAGCCCGCCAGTTTGGAGGAGGGTG 1122
|||||QY 262 CCACAACAGCAGAGTCTCTGAGCTGACACAGGAGCGCGAGTTGGAACCGGGGAGCAG 321
|||||Db 1123 CCACAACAGCAGAGTCTCTGAGCTGACACAGGAGCGCGAGTTGGAACCGGGGAGCAG 1182
|||||QY 322 AGCAGGTGGCCGACGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 381
|||||Db 1183 AGCAGGTGGCCGACGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1242
|||||QY 382 GGCAACATCTACATCTACATCTACATCTACATCTACATCTACATCTACATCTACATCT 441
|||||Db 1243 GGCAACATCTACATCTACATCTACATCTACATCTACATCTACATCTACATCTACATCT 1302
|||||QY 442 CCAGCTACCCCGAACCCTCCATACCCCATTCGCGAAGAGGGGACCCCTGGCCCTCCCGGG 501
|||||Db 1303 CCAGCTACCCCGAACCCTCCATACCCCATTCGCGAAGAGGGGACCCCTGGCCCTCCCGGG 1362
|||||QY 502 CTCTCTACACCCACACAGGAGATGGCAAGCTTGGCACTGATGATGATGATGATGATGATGAT 561
|||||Db 1363 CTCTCTACACCCACACAGGAGATGGCAAGCTTGGCACTGATGATGATGATGATGATGATGAT 1422
|||||QY 562 GGTGCCACACCCCTC 575
|||||Db 1423 GGTGCCACACCCCTC 1436
|||||

RESULT 4

ABL64078

ID ABL64078 standard; DNA; 2136 BP.

XX

AC ABL64078;

XX
DT 15-MAY-2002 (first entry)
XX
DE Breast cancer related gene sequence SEQ ID NO:2415.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
XX 13-DEC-2001.
XX
XX 30-MAY-2001; 2001WO-US10838.
XX
XX 05-JUN-2000; 2000US-209473P.
XX
XX 05-JUN-2000; 2000US-209531P.
XX
XX 18-SEP-2000; 2000US-233133P.
XX
XX 18-SEP-2000; 2000US-233617P.
XX
XX 20-SEP-2000; 2000US-234009P.
XX
XX 20-SEP-2000; 2000US-234034P.
XX
XX 20-SEP-2000; 2000US-234052P.
XX
XX 22-SEP-2000; 2000US-234509P.
XX
XX 22-SEP-2000; 2000US-234567P.
XX
XX 25-SEP-2000; 2000US-234923P.
XX
XX 25-SEP-2000; 2000US-234924P.
XX
XX 25-SEP-2000; 2000US-235077P.
XX
XX 25-SEP-2000; 2000US-235082P.
XX
XX 25-SEP-2000; 2000US-235134P.
XX
XX 26-SEP-2000; 2000US-235280P.
XX
XX 26-SEP-2000; 2000US-235637P.
XX
XX 27-SEP-2000; 2000US-235638P.
XX
XX 27-SEP-2000; 2000US-235711P.
XX
XX 27-SEP-2000; 2000US-235720P.
XX
XX 27-SEP-2000; 2000US-235840P.
XX
XX 27-SEP-2000; 2000US-235863P.
XX
XX 28-SEP-2000; 2000US-236028P.
XX
XX 28-SEP-2000; 2000US-236032P.
XX
XX 28-SEP-2000; 2000US-236033P.
XX
XX 28-SEP-2000; 2000US-236034P.
XX
XX 28-SEP-2000; 2000US-236109P.
XX
XX 28-SEP-2000; 2000US-236111P.
XX
XX 29-SEP-2000; 2000US-236842P.
XX
XX 29-SEP-2000; 2000US-236891P.
XX
XX 02-OCT-2000; 2000US-237172P.
XX
XX 02-OCT-2000; 2000US-237173P.
XX
XX 02-OCT-2000; 2000US-237278P.
XX
XX 02-OCT-2000; 2000US-237294P.
XX
XX 02-OCT-2000; 2000US-237295P.
XX
XX 02-OCT-2000; 2000US-237316P.
XX
XX 03-OCT-2000; 2000US-237425P.
XX
XX 03-OCT-2000; 2000US-237598P.
XX
XX 03-OCT-2000; 2000US-237604P.
XX
XX 03-OCT-2000; 2000US-237606P.
XX
XX 03-OCT-2000; 2000US-237608P.
XX
XX 01-NOV-2000; 2000US-244867P.
XX
XX 01-NOV-2000; 2000US-245084P.
XX
XX (AVAL-) AVALON PHARM.
XX
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
XX Soppet DR, Weaver Z;
XX WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
XX chemical agent to be tested for anti-neoplastic activity, and
XX determining a change in expression of a gene of a signature gene set
XX
XX Claim 1; SEQ ID 2415; 44pp; English.

XX The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX

XX Sequence 2136 BP; 446 A; 706 C; 608 G; 376 T; 0 other;

Query Match 85.9%; Score 494; DB 24; Length 2136;
Best Local Similarity 100.0%; Pred. No. 1.6e-127;
Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 GGATCGCTGCTCAAGAGCGTCCGCGAGGAGGAGGACCAATCCTGTAGCTGGAGCTGG 141
DB 943 GGATCGCTGCTCAAGAGCGTCCGCGAGGAGGAGGACCAATCCTGTAGCTGGAGCTGG 1002
QY 142 GAGCTCCGAAGGCCCATCCATCTTCCCTGACTTGGTACAGCCACTGCTACCCATTCTT 201
DB 1003 GAGCTCCGAAGGCCCATCCATCTTCCCTGACTTGGTACAGCCACTGCTACCCATTCTT 1062
QY 202 GGATGTTTCCCGCAGTATCCACATGGGCTCCCGCAGGCCCGCAGTTTGGAGGCGGGGTG 261
DB 1063 GGATGTTTCCCGCAGTATCCACATGGGCTCCCGCAGGCCCGCAGTTTGGAGGCGGGGTG 1122
QY 262 CCGCAACAGCAGAGTCTCTTGAGCTGACACAGGAGCGCAGTTGGAAACCCGGGAGCAG 321
DB 1123 CCGCAACAGCAGAGTCTCTTGAGCTGACACAGGAGCGCAGTTGGAAACCCGGGAGCAG 1182
QY 322 AGCAGGTGGCCCGCAGGTACCAATGGCATTCATGTCACCGCGGGTCTATGACTATCACT 381
DB 1183 AGCAGGTGGCCCGCAGGTACCAATGGCATTCATGTCACCGCGGGTCTATGACTATCACT 1242
QY 382 GGCACATCTACATCTACATGGACCACTACTGGGGGAGCCACCGGGTCTGGAGACCTC 441
DB 1243 GGCACATCTACATCTACATGGACCACTACTGGGGGAGCCACCGGGTCTGGAGACCTC 1302
QY 442 CCAGCTACCCCGGAACCTCCATACCCCATTCGCCGAAGAGGGGACCCCTGGCCCTCCCGGG 501
DB 1303 CCAGCTACCCCGGAACCTCCATACCCCATTCGCCGAAGAGGGGACCCCTGGCCCTCCCGGG 1362
QY 502 CTCCTACACCCCGCAGGAGAGTGGCAAGGCTTGGCACTTAGCGGAGACAGACACTGT 561
DB 1363 CTCCTACACCCCGCAGGAGAGTGGCAAGGCTTGGCACTTAGCGGAGACAGACACTGT 1422
QY 562 GGTGCCACACCTC 575
DB 1423 GGTGCCACACCTC 1436

RESULT 5
ABL64384
ID ABL64384 standard; DNA; 2136 BP.
XX
XX ABL64384;
XX
XX 15-MAY-2002 (first entry)
XX
XX Stomach cancer related gene sequence SEQ ID NO:2721.
XX
XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;

XX 29-MAY-2001; 2001WO-US17756.
XX
XX
XX 26-MAY-2000; 2000US-207484P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Algate PA, Harlocker SL, Jones R;
XX
XX WPI; 2002-122075/16.
XX
XX Composition for therapy and diagnosis of ovarian cancer comprising
XX polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
XX polypeptide, antibody specific to polypeptide or T cell expressing
XX polypeptide
XX
XX Claim 1; SEQ ID 5601; 489pp; English.
XX
XX The present invention describes a composition (I) comprising: carriers
XX and immunostimulants; and a polypeptide (II) of a ovarian tumor
XX polypeptide encoded by a polynucleotide (III) having a cDNA sequence
XX (S1) from the 10912 nucleotide sequences as given in ABL77023 to
XX ABL87934, (III) encoding (II) having a sequence (S2), a T cell
XX population of (II), or antigen presenting cells that express (II).
XX (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to
XX (S1) can be used for detecting ovarian cancer in a patient's biological
XX sample preferably serum or ovarian tissue. The method comprises
XX contacting a biological sample from a patient with (IV), detecting the
XX amount of polynucleotide hybridising to (IV) and comparing the amount to
XX a predetermined cutoff value and thereby detecting ovarian cancer in the
XX patient, where the amount of polynucleotide hybridising to (IV) is
XX detected preferably by polymerase chain reaction (PCR). (I) comprising
XX (III) and/or (II) is useful for stimulating and/or expanding T cells
XX specific for an ovarian tumor protein comprising contacting T cells
XX with (III) or (II). (III) is useful in design and preparation of
XX ribozyme molecules for inhibiting expression of the tumor polypeptides
XX and proteins in tumor cells; and to isolate a full length gene from a
XX suitable library e.g., a tumor cDNA library using well known
XX techniques.
XX
XX Sequence 281 BP; 58 A; 94 C; 77 G; 52 T; 0 other;
XX
XX Query Match 44.1%; Score 253.8; DB 24; Length 281;
XX Best Local Similarity 97.1%; Pred. No. 6.1e-61;
XX Matches 269; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
XX
QY 193 CCATTTCTGGAGTGTTCCTCCAGTATCCACTGGGCTCCCGAGCC-CAGTTTGGGA 251
Db || |||||
5 CCACATTTCTGGAGTGTTCCTCCAGTATCCACTGGGCTCCCGAGCCCGAGTTTGGGA 64
QY 252 GCGAGGGGTCCGCAACAGCAGAGTCTCTGGACCTGACCGAGCGCGAGTTGGAAC 311
Db |||||
65 GCGAGGGGTCCGCAACAGCAGAGTCTCTGGACCTGACCGAGCGCGAGTTGGAAC 124
QY 312 CGGGAGCAGAGCGAGTGGCCACAGTACCAATGGCATTCATGTACCGCGGGTCTAT 371
Db |||||
125 CGGGAGCAGACTAGGTGGCCCAAGTACCAATGGCATTCATGTACCGCGGGTCTAT 184
QY 372 GACTATCACTGGCAACATCTACATCTACATGGACAGTGGGGGGACCGCGGTCC 431
Db |||||
185 GACTATCACTGGCAACATCTACATCTACATGGACAGTGGGGGGACCGCGGTCA 244
QY 432 TGGAGACCTCCCGAGTACCGCGCAACCTCCATACCCC 468
Db |||||
245 TGGAGACCTCCCGAGTACCGCGCAACCTCCATTCCCC 281
RESULT 7
ID ABL82688/c
XX ABL82688 standard; cDNA; 289 BP.
XX
AC ABL82688;
XX

DT 17-MAY-2002 (first entry)
XX
XX Human ovarian cancer related cDNA clone SEQ ID NO:5666.
XX
XX Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200192581-A2.
XX
XX 06-DEC-2001.
XX
XX 29-MAY-2001; 2001WO-US17756.
XX
XX 26-MAY-2000; 2000US-207484P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Algate PA, Harlocker SL, Jones R;
XX
XX WPI; 2002-122075/16.
XX
XX Composition for therapy and diagnosis of ovarian cancer comprising
XX polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
XX polypeptide, antibody specific to polypeptide or T cell expressing
XX polypeptide
XX
XX Claim 1; SEQ ID 5666; 489pp; English.
XX
XX The present invention describes a composition (I) comprising: carriers
XX and immunostimulants; and a polypeptide (II) of a ovarian tumor
XX polypeptide encoded by a polynucleotide (III) having a cDNA sequence
XX (S1) from the 10912 nucleotide sequences as given in ABL77023 to
XX ABL87934, (III) encoding (II) having a sequence (S2), a T cell
XX population of (II), or antigen presenting cells that express (II).
XX (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to
XX (S1) can be used for detecting ovarian cancer in a patient's biological
XX sample preferably serum or ovarian tissue. The method comprises
XX contacting a biological sample from a patient with (IV), detecting the
XX amount of polynucleotide hybridising to (IV) and comparing the amount to
XX a predetermined cutoff value and thereby detecting ovarian cancer in the
XX patient, where the amount of polynucleotide hybridising to (IV) is
XX detected preferably by polymerase chain reaction (PCR). (I) comprising
XX (III) and/or (II) is useful for stimulating and/or expanding T cells
XX specific for an ovarian tumor protein comprising contacting T cells
XX with (III) or (II). (III) is useful in design and preparation of
XX ribozyme molecules for inhibiting expression of the tumor polypeptides
XX and proteins in tumor cells; and to isolate a full length gene from a
XX suitable library e.g., a tumor cDNA library using well known
XX techniques.
XX
XX Sequence 289 BP; 49 A; 75 C; 93 G; 72 T; 0 other;
XX
XX Query Match 43.7%; Score 251.2; DB 24; Length 289;
XX Best Local Similarity 98.5%; Pred. No. 3.2e-60;
XX Matches 264; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
XX
QY 204 AGATGTTTCCCGAGTATCCACTGGGCTCCCGAGCGCCAGTTTGGAGCGAGGTGCC 263
Db |||||
289 AGATGTTTCCCGAGTATCCACTGGGCTCCCGAGCGCCAGTTTGGAGCGAGGTGCC 230
QY 264 GCAACAGCAGA-GTCTCTGGACCTGACACAGGAGCCCGAGTTGGAAACCCGGGAGCAGA 322
Db |||||
229 GCAACAGCAGATGTCTCTGGACCTGACACAGGAGCCCGGTGGAAACCCGGGAGCAGA 170
QY 323 GCCAGGTGCCCGACGCTACCATGTCATGTACCGCGGGTCTATGACTATCACTG 382
Db |||||
169 GCCAGGTGCCCGACCGTACCAATGGCATTCATGTACCGCGGGTCTATGACTATCACTG 110
QY 383 GCAACATCTACATCTACATGGACAGTACTGGGGGGACCGCGGGTCTCTGAGACCTCC 442
Db |||||
109 GCAACATCTACATCTACATGGACAGTACTGGGGGGACCGCGGGTCTCTGAGACCTCC 50

QY 443 CAGCTACCCCGGACCTCCATACCCCAT 470
 |||||
 DB 49 CAGCTACCCCGGACCTCCATTCGCCCT 22

RESULT 8

ABL81903
 ID ABL81903 standard; cDNA; 481 BP.

XX AC ABL81903;
 XX DT 17-MAY-2002 (first entry)
 XX DE Human ovarian cancer related cDNA clone SEQ ID NO:4881.

XX KW Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
 XX OS Homo sapiens.

XX PN WO200192581-A2.

XX PD 06-DEC-2001.

XX PF 29-MAY-2001; 2001WO-US17756.
 XX PR 26-MAY-2000; 2000US-207484P.

XX PA (CORI-) CORIXA CORP.

XX PI Algate PA, Harlocker SL, Jones R;
 XX DR WPI; 2002-122075/16.

XX CC Composition for therapy and diagnosis of ovarian cancer comprising
 PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
 PT polypeptide, antibody specific to polypeptide or T cell expressing
 PT polypeptide -

XX PS Claim 1; SEQ ID 4881; 489pp; English.

XX CC The present invention describes a composition (I) comprising: carriers
 CC and immunostimulants; and a polypeptide (II) of a ovarian tumour
 CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence
 CC (SI) from the 10912 nucleotide sequences as given in ABL77023 to
 CC ABL87934, (III) encoding (II) having a sequence (S2), a T cell
 CC population of (II), or antigen presenting cells that express (II).
 CC (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to
 CC (SI) can be used for detecting ovarian cancer in a patient's biological
 CC sample preferably serum or ovarian tissue. The method comprises
 CC contacting a biological sample from a patient with (IV), detecting the
 CC amount of polynucleotide hybridising to (IV) and comparing the amount to
 CC a predetermined cutoff value and thereby detecting ovarian cancer in the
 CC patient, where the amount of polynucleotide hybridising to (IV) is
 CC detected preferably by polymerase chain reaction (PCR). (I) comprising
 CC (III) and/or (II) is useful for stimulating and/or expanding T cells
 CC specific for an ovarian tumour protein comprising contacting T cells
 CC with (III) or (II). (III) is useful in design and preparation of
 CC ribozyme molecules for inhibiting expression of the tumour polypeptides
 CC and proteins in tumour cells; and to isolate a full length gene from a
 CC suitable library e.g., a tumour cDNA library using well known
 CC techniques.

XX SQ Sequence 481 BP; 113 A; 161 C; 123 G; 84 T; 0 other;

Query Match 10.6%; Score 61.2; DB 24; Length 481;
 Best Local Similarity 95.5%; Pred. NO. 2.9e-07;
 Matches 63; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTGTGGAGGAGCTCCAGGCACTGCCGACACCACTGCAGAAATCCATTAGAG 60
 |||||

DB 416 CTGTGGAGGAGCTCCAGGCACTGCACAGTCCACACCACTGCAGAAATCCATTAGAG 475

QY 61 CCACTG 66

DB 476 CCACTG 481
 |||||

RESULT 9

AAV72224
 ID AAV72224 standard; DNA; 333 BP.

XX AC AAV72224;

XX DT 07-SEP-1999 (first entry)

XX DE Human anti-GPIIb/IIIa antibody light chain DNA from phagemid PD67.

XX KW Antibody; GPIIb/IIIa; human; auto-antibody; anti-idiotypic; diagnosis;
 KW blood platelet membrane protein; predisposition; prevention; treatment;
 KW autoimmune thrombocytopenic purpura; AITP; fibrinogen binding; thrombi;
 KW thrombocyte; cardiac infarction; pulmonary embolism; light chain; ds.

XX OS Homo sapiens.

XX PH Key
 XX FT misc_feature
 FT /tag= a
 FT /note= "Framework region 1 (FR1)"
 FT 61..99

FT misc_feature
 FT /tag= b
 FT /note= "complementarity determining region 1 (CDR1)"
 FT 100..144

FT misc_feature
 FT /tag= c
 FT /note= "Framework region 2 (FR2)"
 FT 145..165

FT misc_feature
 FT /tag= d
 FT /note= "complementarity determining region 2 (CDR2)"
 FT 166..261

FT misc_feature
 FT /tag= e
 FT /note= "Framework region 3 (FR3)"
 FT 262..294

FT misc_feature
 FT /tag= f
 FT /note= "complementarity determining region 3 (CDR3)"
 FT 295..333

FT misc_feature
 FT /tag= g
 FT /note= "Framework region 4 (FR4)"
 FT 334..399

XX WO9855619-A1.

XX PD 10-DEC-1998.

XX PF 05-JUN-1998; 98WO-EP03397.

XX PR 08-MAY-1998; 98DE-1020663.

XX PR 06-JUN-1997; 97DE-1023904.

XX PR 12-DEC-1997; 97DE-1055227.

XX PA (ASAT-) ASAT AG APPLIED SCI & TECHNOLOGY.

XX PI Berchtold P, Escher RFA;

XX DR WPI; 1999-105496/09.

XX DR P-PSDB; AAW90279.

XX PT Nucleic acid encoding human autoantibodies against platelet
 PT glycoprotein IIb/IIIa - used for diagnosis, treatment and prevention
 PT of autoimmune thrombocytopenic purpura and for modulation of
 PT fibrinogen binding

XX PS Disclosure; Page 49-50; 93pp; German.

XX CC This invention describes novel nucleic acid fragments that encode human
 CC auto-antibodies and anti-idiotypic antibodies against blood platelet
 CC membrane protein, GPIIb/IIIa. The products of the invention are used
 CC for diagnosis (including monitoring and determining predisposition),
 CC prevention and treatment of autoimmune thrombocytopenic purpura (AITP)

CC and also for modulating binding of fibrinogen to thrombocytes
 CC (particularly to dissolve thrombi and/or prevent their formation, e.g.
 CC in cases of cardiac infarction or pulmonary embolism). Unlike murine
 CC antibodies, human antibodies (hAb) do not induce adverse side effects
 CC and persist for longer in vivo than small peptides.
 XX
 SQ Sequence 333 BP; 59 A; 109 C; 93 G; 72 T; 0 other;
 Query Match 6.3%; Score 36.4; DB 20; Length 333;
 Best Local Similarity 49.0%; Pred. No. 2.1;
 Matches 97; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
 QY 5 TGGAGGAGCTCCAGGAGCTCCGACGACCAACCTGCAAAATCCATTAGAGCCAC 64
 Db 63 TGGAGGAGCTCCAGGAGCTCCGACGACCAACCTGCTAGCTGGTATCACCAGGTCCCGAGCAC 122
 QY 65 TGCCCCCAGAGATGTCAGGATCGCTCAAGAGCGCTCCGAGGGGAGAGGCCCAATC 124
 Db 123 GGCCCCCAACTCTCATCTTTGTAGTATCATGACGGCCCTCAGGGGTCCCTGACCGATT 182
 QY 125 CTGTAGCTGGAAGCTGGGAGCCTCCGAAGGCCATCATCTTCCCTGACTTGGTACAGC 184
 Db 183 CTCGTGGCTCAAGTCGGGACCTCCGCTCCCTGGCCATCGTGGGTCCCAATCTGGGA 242
 QY 185 CACTGTACCAATTCG 202
 Db 243 TGCTGGTGAATTAATG 260
 RESULT 10
 AAH07782/c
 ID AAH07782 standard; cDNA; 788 BP.
 XX
 AC AAH07782;
 XX
 DT 26-JUN-2001 (first entry)
 DE Human cDNA clone (5'-primer) SEQ ID NO:4617.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 PS Claim 1; SEQ ID 4617; 2537pp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of

CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 788 BP; 179 A; 204 C; 249 G; 153 T; 3 other;
 Query Match 6.3%; Score 36.4; DB 22; Length 788;
 Best Local Similarity 51.2%; Pred. No. 2.7;
 Matches 85; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
 QY 404 GACCACTACTTGGGGGACCAACCGGCTCTGGAGACCTCCAGCTACCCCGAACCTCCAT 463
 Db 414 GGCAGGCGCTGGGAGATGCACAGGCGCCCGGAGACGTGTCTGCCCTCCAGAGCGCTGCC 355
 QY 464 ACCCAATTCGGAAGAGGGGACCTCCGCGCTCCCGGGTCTCTACACCCACCCAGGAAG 523
 Db 354 TTCTCATCTATGAGGAGCGGGGACGATGCCAAAGAGGGCCCTCATCGCCATGGCAGGAGG 295
 QY 524 ATGGCAAGGCTTGGCACCTAGCGAGACAGACGACTGTGGTGCCAC 569
 Db 294 ATTTCCGCCCTTCTTCCGAGGGCGGTGGTGGATGCAGATGCCAC 249
 RESULT 11
 AA193910/c
 ID AA193910 standard; cDNA; 2497 BP.
 XX
 AC AA193910;
 XX
 DT 13-NOV-2001 (first entry)
 DE Human stomach cancer expressed polynucleotide SEQ ID NO 134.
 XX
 KW Human; stomach cancer; marker; screening; micro-metastasis;
 KW peritoneal dissemination; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200109317-A1.
 XX
 PD 08-FEB-2001.
 XX
 PF 28-JUL-2000; 2000WO-JP05063.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 18-OCT-1999; 99US-0159590.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 17-FEB-2000; 2000US-0183322.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T, Aburatani H;
 PI Kodama T, Midorikawa Y;
 XX

PR 29-NOV-1999; 99US-0167824.
PR 15-FEB-2000; 2000US-0182711.
XX (ALPH-) ALPHAGENE INC.
XX Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;
XX WPI; 2000-638211/61.
XX P-PSDB; AAB34712.
XX Novel proteins and polypeptides useful for the treatment of e.g
XX multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis,
XX cancer, Alzheimer's disease, Parkinson's disease, stroke, anemia and
XX ulcers
XX
XX Claim 60; Page 419-420; 493pp; English.
XX
XX This invention relates to 59 human secreted proteins and the nucleotide
XX sequences encoding them. Sequences AAC59788-C59846 and AAB34687-B34745
XX represent the proteins and their encoding nucleotide sequences, and
XX sequences AAB34746-B34771 represent fragments of the proteins. Probes
XX for the DNA sequences are represented by sequences AAC59847-C59596. The
XX proteins exhibit neuroprotective, dermatological, immunosuppressive,
XX antiinflammatory, antianemic, nootropic, antiparkinsonian,
XX cerebroprotective, haemostatic, vulnerary, cytostatic, antipsoriatic,
XX antibacterial, virucide, and fungicide activity. The proteins and
XX nucleotide sequences are useful as nutritional sources or supplements
XX and in research. The proteins are useful for treating immune deficiency
XX and disorders, which may be genetic or resulting from infections,
XX autoimmune disorders such as multiple sclerosis, systemic lupus
XX erythematosus, rheumatoid arthritis, and for treating myeloid or lymphoid
XX cell deficiencies such as anaemias by regulating haematopoiesis. The
XX proteins are also useful in compositions for bone, cartilage, tendon,
XX ligament and/or nerve tissue growth or regeneration, for wound healing,
XX tissue repair and replacement and in the treatment of wounds, incisions
XX and ulcers. Other uses include in the treatment of central and
XX peripheral nervous system and neuropathies such as Alzheimer's and
XX Parkinson's diseases and Shy-Drager syndrome, and mechanical and
XX traumatic disorders, such as spinal cord disorders, head trauma and
XX stroke. The proteins may also be used as a contraceptive, and for
XX treating coagulation disorders such as haemophilias. The protein and
XX nucleotide sequences with cadherin activity are useful for treating
XX cancer. Other uses for the protein include for inhibiting the growth,
XX infection or function of, or killing, infectious agents such as bacteria,
XX virus, fungi and other parasites, for effecting bodily characteristics
XX such as height, weight, hair colour, effecting biorhythms or cardiac
XX cycles or rhythms, effecting metabolism, catabolism, anabolism,
XX processing, utilization, storage or elimination of dietary fat, lipid,
XX protein, carbohydrate, vitamins, minerals, cofactors, effecting
XX behavioural characteristics, providing analgesic effects and for treating
XX hyperproliferative disorders such as psoriasis.
XX
XX Sequence 3290 BP; 729 A; 816 C; 971 G; 774 T; 0 other;
XX
XX Query Match 6.3%; Score 36.4; DB 21; Length 3290;
XX Best Local Similarity 51.2%; Pred. No. 4.4;
XX Matches 85; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
XX
XX QY 404 GACCAGTACTGGGGGACCCGCTCTGGAGACCTCCAGCTACCCCGAACCTCCAT 463
XX Db 1183 GCCAGCGCTGGGAGATGCACAGGGCCCGGAGAGCTGTGCTCCACGACGCTGCC 1124
XX
XX QY 464 ACCCCATTCCGAAGAGGGGACCTGGCCCTCCCGGGCTCTACACCCACCAGGAAG 523
XX Db 1123 TTCTCATCCATGAGACGGGACGATGCCAAGAGAGGGCCCTATCGCCATGGGAGG 1064
XX
XX QY 524 ATGCAAGGCTTGACCTAGCGGAGACAGAGACTGTGGTGCAC 569
XX Db 1063 ATTTCCGCCCTTCTCCGAGGGCCGTGGTGGATGCAGATGCCAC 1018
XX
XX RESULT 14
XX AAS40920/c

AAAS40920 standard; cDNA; 3298 BP.
AAS40920;
17-DEC-2001 (first entry)
cDNA encoding novel human enzyme polypeptide #136.
Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
ligase; hyperproliferative disorder; immunodeficiency disorder;
autoimmune disorder; neurological disorder; metabolic disorder;
inflammatory disorder; cardiovascular disorder; reproductive disorder;
blood-related disorder; infectious disorder; gene therapy; cytostatic;
anti arthritic; nephrotropic; anticoagulant; ss.
Homo sapiens.
WO200155301-A2.
02-AUG-2001.
17-JAN-2001; 2001WO-US01239.
31-JAN-2000; 2000US-0179065.
04-FEB-2000; 2000US-0180628.
24-FEB-2000; 2000US-0184664.
02-MAR-2000; 2000US-0186350.
16-MAR-2000; 2000US-0189874.
17-MAR-2000; 2000US-0190076.
18-APR-2000; 2000US-0198123.
19-MAY-2000; 2000US-0205515.
07-JUN-2000; 2000US-0209467.
28-JUN-2000; 2000US-0214886.
30-JUN-2000; 2000US-0215135.
07-JUL-2000; 2000US-0216647.
07-JUL-2000; 2000US-0216880.
11-JUL-2000; 2000US-0217487.
11-JUL-2000; 2000US-0217496.
14-JUL-2000; 2000US-0218290.
26-JUL-2000; 2000US-0220963.
26-JUL-2000; 2000US-0220964.
14-AUG-2000; 2000US-0224518.
14-AUG-2000; 2000US-0224519.
14-AUG-2000; 2000US-0225213.
14-AUG-2000; 2000US-0225214.
14-AUG-2000; 2000US-0225266.
14-AUG-2000; 2000US-0225267.
14-AUG-2000; 2000US-0225268.
14-AUG-2000; 2000US-0225270.
14-AUG-2000; 2000US-0225447.
14-AUG-2000; 2000US-0225757.
14-AUG-2000; 2000US-0225758.
14-AUG-2000; 2000US-0225759.
18-AUG-2000; 2000US-0226279.
22-AUG-2000; 2000US-0226681.
22-AUG-2000; 2000US-0226868.
22-AUG-2000; 2000US-0227182.
23-AUG-2000; 2000US-0227009.
30-AUG-2000; 2000US-0228924.
01-SEP-2000; 2000US-0229287.
01-SEP-2000; 2000US-0229343.
01-SEP-2000; 2000US-0229344.
01-SEP-2000; 2000US-0229345.
05-SEP-2000; 2000US-0229509.
05-SEP-2000; 2000US-0229513.
06-SEP-2000; 2000US-0230437.
06-SEP-2000; 2000US-0230438.
08-SEP-2000; 2000US-0231242.
08-SEP-2000; 2000US-0231243.
08-SEP-2000; 2000US-0231244.
08-SEP-2000; 2000US-0231413.
08-SEP-2000; 2000US-0231414.
08-SEP-2000; 2000US-0232080.

PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 14-SEP-2000; 2000US-0233066.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.

PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PA (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-465566/50.
XX P-PSDB; AAU23050.
XX Novel polypeptides and polynucleotides useful for diagnosing,
PT preventing, treating neural, immune system, muscular, reproductive,
PT pulmonary, cardiovascular, renal, proliferative disorders and cancerous
PT diseases -
XX Claim 4; SEQ ID No 146; 1180pp; English.
XX The present invention relates to the isolation of novel human enzyme
CC polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences
CC encoding them. The enzyme polypeptides of the invention may comprise the
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
CC isomerases or ligases. The sequences of the invention are useful in the
CC diagnosis, treatment, prevention and/or prognosis of a wide range of
CC disorders including hyperproliferative disorders (e.g. cancer),
CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders
CC (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),
CC metabolic disorders (e.g. phenylketonuria), inflammatory disorders
CC (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),
CC blood-related disorders (e.g. haemophilia), reproductive disorders
CC (e.g. infertility) and infectious disorders (e.g. influenza). The
CC polynucleotides of the invention can also be used in gene therapy.
CC AAU20785-AAU21684 represent cDNA sequences encoding for the novel human
CC enzyme polypeptides of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 3298 BP; 744 A; 817 C; 968 G; 769 T; 0 other;
SQ
Query Match 6.3%; Score 36.4; DB 22; Length 3298;
Best Local Similarity 51.2%; Pred. No. 4.4;
Matches 85; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
QY 404 GACCAGTACTGGGGGACACCGGGTCTCTGGAGACTCCAGCTACCCCGAACCTCCAT 463
Db 1173 GCCCAGGCTGGAGATGCACAGGGCCCGGAGACGTTCTGCTCCACGAGCTGCC 1114
QY 464 ACCCCATTCCCGAAGAGGGGACCCCTGGCCCTCCCGGGCTCTCTACACCCACAGGAAG 523
Db 1113 TTCTCATCATGAGGACGGGACGATGCCAAAGAGGGCTCATGCCATGGCAGGAGG 1054
QY 524 ATGGCAAGGCTTGGCACCTAGCGGAGACAGCAGTGTGGTCCAC 569
Db 1053 ATTTCGGCCCTTCTTCGAGGGCCGCTGGTGGATGCAGATGCCAC 1008
RESULT 15
ABV38197
ID ABV38197 standard; cDNA; 450 BP.
XX AC
XX ABV38197;
XX

Search completed: April 15, 2003, 15:16:36
Job time : 112.083 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:09:45 ; Search time 664.156 Seconds
(without alignments)
14021.414 Million cell updates/sec

Title: US-09-917-372-2_COPY_804_1378

Perfect score: 575

Sequence: 1 ctggtggaggcagctccagg.....cactgtggtgccacaccctc 575

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	494	85.9	814	9	AU118203
2	494	85.9	868	10	BE275240
3	494	85.9	895	14	BQ957322
4	494	85.9	920	12	BG036632
5	494	85.9	954	13	BI821009
6	494	85.9	968	14	BQ898015

7	482	83.8	714	13	BI829057
8	482	83.8	901	12	BE740143
9	481.4	83.7	954	14	BQ930450
10	480.4	83.5	900	13	BI769498
11	452.8	78.7	749	12	BE789430
12	445	77.4	735	12	BG331666
13	444.8	77.4	606	13	BI905922
14	437.2	76.0	700	12	BG576875
15	432	75.1	598	12	BG403820
16	425.4	74.0	427	10	AW630662
17	423	73.6	801	12	BF792867
18	417	72.5	458	14	BM766807
19	417	72.5	537	14	BM766352
20	417	72.5	572	14	BM767287
21	417	72.5	666	14	BM767318
22	416	72.3	501	14	BM766833
23	415.4	72.2	634	14	BM746911
24	412	71.7	598	14	BM723050
25	409.4	71.2	579	10	BE207591
26	381	66.3	640	12	BG491235
27	376	65.4	506	14	BM855039
28	375	65.2	808	13	BI818411
29	374.4	65.1	466	14	BQ082995
30	370.8	64.5	426	14	BM856797
31	362.8	63.1	416	14	BQ083004
32	361.2	62.8	418	14	BM855410
33	350.2	60.9	942	13	BI818552
34	347.6	60.5	721	9	AI336202
35	344.8	60.0	360	14	BM766382
36	337	58.6	706	12	BG698325
37	332.2	57.8	633	9	AI870838
38	326	56.7	728	13	BI765703
39	317.2	55.2	530	12	BG610053
40	315.4	54.9	386	9	AA099514
41	314	54.6	882	14	BQ691803
42	306.8	53.4	518	13	BI338842
43	282	49.0	598	14	BQ305454
44	278.8	48.5	665	13	BI554834
45	278.8	48.5	934	14	BQ928227

ALIGNMENTS

RESULT 1	AU118203	814 bp	linear	EST 01-AUG-2002
LOCUS	AU118203	HEMBA1	Homo sapiens	cdna clone HEMBA1003089 5', mRNA
DEFINITION	AU118203	sequence.		
ACCESSION	AU118203			
VERSION	AU118203.1	GI:10933231		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and Isogai, T.			
TITLE	1 (bases 1 to 814)			
JOURNAL	Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and Isogai, T.			
COMMENT	Unpublished (2000)			
	Contact: Takao Isogai			
	Genomics Laboratory			
	Helix Research Institute			
	1532-3 Yana, Kisarazu, Chiba 292-0812, Japan			
	Tel: 81-438-52-3975			
	Fax: 81-438-52-3986			
	Email: genomics@hri.co.jp			
	HRI human cdna project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cdna library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.			

```

FEATURES
  source      Location/Qualifiers
  1..814
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="HEMBA1003089"
    /clone_lib="HEMBA1"
    /tissue_type="whole embryo, mainly head"
    /dev_stage="embryo, 10 weeks"
    /note="Vector: pME18SFL3"
  BASE COUNT      173 a      255 c      232 g      151 t
  ORIGIN
    Query Match      85.9%; Score 494; DB 9; Length 814;
    Best Local Similarity 100.0%; Pred. No. 1.3e-111;
    Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 82 GGATCGCTCTCAAGAGGGCTCCGAGGAGGAGGACCAATCCCTAGCTGGAAGCTGG 141
  Db 98 GGATCGCTCTCAAGAGGGCTCCGAGGAGGAGGACCAATCCCTAGCTGGAAGCTGG 157
  QY 142 GAGCCTCCGAAGGCCCATCATCTACTTCCCTGACTTGGTACAGCCACTGCTACCCATTCT 201
  Db 158 GAGCCTCCGAAGGCCCATCATCTACTTCCCTGACTTGGTACAGCCACTGCTACCCATTCT 217
  QY 202 GGAGATGTTTCCCACTATCCACTGGGCTCCCGGAGGCCCTTGGAGGAGGGGTG 261
  Db 218 GGAGATGTTTCCCACTATCCACTGGGCTCCCGGAGGCCCTTGGAGGAGGGGTG 277
  QY 262 CGCAACAGCAGAGTCTCTGACCTGACCGAGGAGCGCAGTTGGAACCCGGGAGCAG 321
  Db 278 CGCAACAGCAGAGTCTCTGACCTGACCGAGGAGCGCAGTTGGAACCCGGGAGCAG 337
  QY 322 AGCCAGGTGGCCCGTACCAATGGCATTCATGTCACCGGGGTCTATGACTATCACT 381
  Db 338 AGCCAGGTGGCCCGTACCAATGGCATTCATGTCACCGGGGTCTATGACTATCACT 397
  QY 382 GGCACATCTACATCTACATGACCATGAGTGGGGGAGCACCGGGTCTGGAGACCTC 441
  Db 398 GGCACATCTACATCTACATGACCATGAGTGGGGGAGCACCGGGTCTGGAGACCTC 457
  QY 442 CCAGTACCCCGAAGCTCCATACCCCATTCGGAAGAGGGGACCCCTGGCCCTCCGGG 501
  Db 458 CCAGTACCCCGAAGCTCCATACCCCATTCGGAAGAGGGGACCCCTGGCCCTCCGGG 517
  QY 502 CTCTCTACACCCCGAAGATGGCAAGGCTTGGCACCTAGCGGAGACAGACACTGT 561
  Db 518 CTCTCTACACCCCGAAGATGGCAAGGCTTGGCACCTAGCGGAGACAGACACTGT 577
  QY 562 GGTGCCACACCCCTC 575
  Db 578 GGTGCCACACCCCTC 591
  RESULT 2
  BE275240
  LOCUS      601122062P1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3346311 5',
  DEFINITION      mRNA sequence.
  ACCESSION      BE275240
  VERSION        BE275240.1 GI:9150193
  KEYWORDS       EST.
  SOURCE         human.
  ORGANISM       Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
    1 (bases 1 to 868)
    NIH-MGC http://mgi.nci.nih.gov/.
    National Institutes of Health, Mammalian Gene Collection (MGC)
    Unpublished (1999)
    Contact: Robert Strausberg, Ph.D.
    Email: cgapbs-remail.nih.gov
    Tissue Procurement: ATCC/DCTD/DPF
    cDNA Library Preparation: Ling Hong/Rubin Laboratory

```

```

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LICM135 row: k column: 16
High quality sequence stop: 746.
FEATURES
  Location/Qualifiers
  1..868
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:3346311"
    /clone_lib="NIH_MGC_20"
    /tissue_type="melanotic melanoma"
    /lab_host="DH10B (phage-resistant)"
    /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
    EcoRI; cDNA made by oligo-dr priming. Directionally
    cloned into EcoRI/XhoI sites using the following 5'
    adaptor: GGCACGAG(G). Size-selected >500bp for average
    insert size 1.8kb. Library constructed by Ling Hong in
    the laboratory of Gerald M. Rubin (University of
    California, Berkeley) using ZAP-cDNA synthesis kit
    (Stratagene) and Superscript II RT (Life Technologies)."
  BASE COUNT      170 a      286 c      253 g      159 t
  ORIGIN
    Query Match      85.9%; Score 494; DB 10; Length 868;
    Best Local Similarity 100.0%; Pred. No. 1.3e-111;
    Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 82 GGATCGCTCTCAAGAGGGCTCCGAGGAGGAGGACCAATCCCTAGCTGGAAGCTGG 141
  Db 138 GGATCGCTCTCAAGAGGGCTCCGAGGAGGAGGACCAATCCCTAGCTGGAAGCTGG 197
  QY 142 GAGCCTCCGAAGGCCCATCATCTACTTCCCTGACTTGGTACAGCCACTGCTACCCATTCT 201
  Db 198 GAGCCTCCGAAGGCCCATCATCTACTTCCCTGACTTGGTACAGCCACTGCTACCCATTCT 257
  QY 202 GGAGATGTTTCCCACTATCCACTGGGCTCCCGGAGGCCCTTGGAGGAGGGGTG 261
  Db 258 GGAGATGTTTCCCACTATCCACTGGGCTCCCGGAGGCCCTTGGAGGAGGGGTG 317
  QY 262 CGCAACAGCAGAGTCTCTGACCTGACCGAGGAGCGCAGTTGGAACCCGGGAGCAG 321
  Db 318 CGCAACAGCAGAGTCTCTGACCTGACCGAGGAGCGCAGTTGGAACCCGGGAGCAG 377
  QY 322 AGCCAGGTGGCCCGTACCAATGGCATTCATGTCACCGGGGTCTATGACTATCACT 381
  Db 378 AGCCAGGTGGCCCGTACCAATGGCATTCATGTCACCGGGGTCTATGACTATCACT 437
  QY 382 GGCACATCTACATCTACATGACCATGAGTGGGGGAGCACCGGGTCTCTGGAGACCTC 441
  Db 438 GGCACATCTACATCTACATGACCATGAGTGGGGGAGCACCGGGTCTCTGGAGACCTC 497
  QY 442 CCAGTACCCCGAAGCTCCATACCCCATTCGGAAGAGGGGACCCCTGGCCCTCCGGG 501
  Db 498 CCAGTACCCCGAAGCTCCATACCCCATTCGGAAGAGGGGACCCCTGGCCCTCCGGG 557
  QY 502 CTCTCTACACCCCGAAGATGGCAAGGCTTGGCACCTAGCGGAGACAGACACTGT 561
  Db 558 CTCTCTACACCCCGAAGATGGCAAGGCTTGGCACCTAGCGGAGACAGACACTGT 617
  QY 562 GGTGCCACACCCCTC 575
  Db 618 GGTGCCACACCCCTC 631

```

```

RESULT 3
BQ957322
LOCUS
DEFINITION      BQ957322 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6384424
ACCESSION      BQ957322
VERSION        BQ957322.1 GI:22372800

```



```

QY 382 GCAACATCTACATCTACATGAGCAGTACTGGGGGACCAACCGGGTCTCTGGAGACCTC 441
|||||
Db 352 GCAACATCTACATCTACATGAGCAGTACTGGGGGACCAACCGGGTCTCTGGAGACCTC 411
|||||
QY 442 CCAGTACCCCGAGACCTCTCATACCCCATTCCTCCGAGAGGGGACCTGGCCCTCCCGGG 501
|||||
Db 412 CCAGTACCCCGAGACCTCTCATACCCCATTCCTCCGAGAGGGGACCTGGCCCTCCCGGG 471
|||||
QY 502 CTCTCTACACCCACCAAGAGATGGCAAGCTTGGCACCTAGCCGAGACAGACAGACTGT 561
|||||
Db 472 CTCTCTACACCCACCAAGAGATGGCAAGCTTGGCACCTAGCCGAGACAGACAGACTGT 531
|||||
QY 562 GGTGCCACACCCCTC 575
|||||
Db 532 GGTGCCACACCCCTC 545
|||||

RESULT 5
BI821009
LOCUS
DEFINITION
603035664F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5176935 5',
mRNA sequence.
ACCESSION
BI821009
VERSION
BI821009.1 GI:15932559
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 954)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Tissue Procurement: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1440 row: 0 column: 16
High quality sequence stop: 758.
FEATURES
source
1..954
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5176935"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/notes="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."
BASE COUNT 212 a 313 c 281 g 148 t
ORIGIN

Query Match 85.9%; Score 494; DB 13; Length 954;
Best Local Similarity 100.0%; Pred. No. 1.3e-111;
Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 GGATCGCTGCTCAGAGCGCTCCGAGGAGAGGGACCAATCTGTAGCTGGAAGCTGG 141
|||||
Db 14 GGATCGCTGCTCAGAGCGCTCCGAGGAGAGGGACCAATCTGTAGCTGGAAGCTGG 73
|||||

```

```

QY 142 GAGCTCCGAAGGCCCATCTCCCTGACTTGTGTACAGCCACTGCTACCCATTCT 201
|||||
Db 74 GAGCTCCGAAGGCCCATCTCCCTGACTTGTGTACAGCCACTGCTACCCATTCT 133
|||||
QY 202 GGAGATGTTTCCCGAGTATCCACTGGGCTCCCGCAGCCCCAGTTTGGAGGCGAGGGTG 261
|||||
Db 134 GGAGATGTTTCCCGAGTATCCACTGGGCTCCCGCAGCCCCAGTTTGGAGGCGAGGGTG 193
|||||
QY 262 CCGCAACAGCAGAGTCTCTCTGGACCTGACACGAGGAGCGCAGTTGGAAACCCGGGAGCAG 321
|||||
Db 194 CCGCAACAGCAGAGTCTCTCTGGACCTGACACGAGGAGCGCAGTTGGAAACCCGGGAGCAG 253
|||||
QY 322 AGCCAGGTGGCCCGGCTACCAATGGCATTCATCTACCGGGGGTCTATGACTATCACT 381
|||||
Db 254 AGCCAGGTGGCCCGGCTACCAATGGCATTCATCTACCGGGGGTCTATGACTATCACT 313
|||||
QY 382 GGCAACATCTACATCTACAATGGACCACTACTGGGGGACCAACCGGGTCTCTGGAGACCTC 441
|||||
Db 314 GGCAACATCTACATCTACAATGGACCACTACTGGGGGACCAACCGGGTCTCTGGAGACCTC 373
|||||
QY 442 CCAGTACCCCGAGACCTCTCATACCCCATTCCTCCGAGAGGGGACCTGGCCCTCCCGGG 501
|||||
Db 374 CCAGTACCCCGAGACCTCTCATACCCCATTCCTCCGAGAGGGGACCTGGCCCTCCCGGG 433
|||||
QY 502 CTCTCTACACCCACCAAGAGATGGCAAGCTTGGCACCTAGCCGAGACAGACAGACTGT 561
|||||
Db 434 CTCTCTACACCCACCAAGAGATGGCAAGCTTGGCACCTAGCCGAGACAGACAGACTGT 493
|||||
QY 562 GGTGCCACACCCCTC 575
|||||
Db 494 GGTGCCACACCCCTC 507
|||||

RESULT 6
BO898015
LOCUS
DEFINITION
BO898015 968 bp mRNA linear EST 16-AUG-2002
AGENCOURT_8125942 Lupski_dorsal_root_ganglion Homo sapiens cDNA
clone IMAGE:6177574 5', mRNA sequence.
ACCESSION
BO898015
VERSION
BO898015.1 GI:22290029
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 968)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13555 row: 1 column: 23
High quality sequence stop: 615.
FEATURES
source
1..968
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6177574"
/clone_lib="Lupski_dorsal_root_ganglion"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6 (Life Technologies); Site 1:
NotI; Site 2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:

```


5'-TCGACCCACGCGTCG-3' and
5'-GACATGTTAGTCGCGAGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

BASE COUNT 219 a 320 c 260 g 168 t 1 others
ORIGIN

Query Match 85.9%; Score 494; DB 14; Length 968;
Best Local Similarity 100.0%; Pred. No. 1.3e-111; Indels 0; Gaps 0;
Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 GGATCGCTCTCAAGAGCGTCGCGAGGAGGAGGACCAATCTGTAGCTGGAAGCTGG 141
|||||
Db 266 GGATCGCTCTCAAGAGCGTCGCGAGGAGGAGGACCAATCTGTAGCTGGAAGCTGG 325
|||||
QY 142 GAGCTCCGAAGCGCATCATCTCTGACTTGGTACAGCCACTGTACCCATTTCT 201
|||||
Db 326 GAGCTCCGAAGCGCATCATCTCTGACTTGGTACAGCCACTGTACCCATTTCT 385
|||||
QY 202 GGAGATGTTTCCCGAGTATCCACTGGCTCCCGAGCGCCAGTTTGGAGCAGGGGTG 261
|||||
Db 386 GGAGATGTTTCCCGAGTATCCACTGGCTCCCGAGCGCCAGTTTGGAGCAGGGGTG 445
|||||
QY 262 CCGCAACAGCAGAGTCTCTGACCTGACAGGAGCGCAGTTGGAACCGGGGAGCAG 321
|||||
Db 446 CCGCAACAGCAGAGTCTCTGACCTGACAGGAGCGCAGTTGGAACCGGGGAGCAG 505
|||||
QY 322 AGCCAGGTGGCCAGGTACCAATGGCAATGATGACGGGGGTCTATGACTATCACT 381
|||||
Db 506 AGCCAGGTGGCCAGGTACCAATGGCAATGATGACGGGGGTCTATGACTATCACT 565
|||||
QY 382 GGCACATCTACATCTACATGGACGACTGAGGGGAGCAGCGGGTCTTGAGACCTC 441
|||||
Db 566 GGCACATCTACATCTACATGGACGACTGAGGGGAGCAGCGGGTCTTGAGACCTC 625
|||||
QY 442 CCAGCTACCCCGAAGCTTCATACCCATCCCGAAGAGGGGACCCCTGGCCCTCCCGGG 501
|||||
Db 626 CCAGCTACCCCGAAGCTTCATACCCATCCCGAAGAGGGGACCCCTGGCCCTCCCGGG 685
|||||
QY 502 CTCTCTACACCCACAGGAAGATGGCAAGCTTGGCAGCTAGCGGAGACAGCACTGT 561
|||||
Db 686 CTCTCTACACCCACAGGAAGATGGCAAGCTTGGCAGCTAGCGGAGACAGCACTGT 745
|||||
QY 562 GGTGCCACACCTC 575
|||||
Db 746 GGTGCCACACCTC 759

RESULT 7
BI829057

LOCUS 603079147F1 NIH_MGC_119 Homo sapiens cdna clone IMAGE:5171122 5',
mRNA sequence.

DEFINITION BI829057

ACCESSION BI829057.1 GI:15940607

VERSION EST.

KEYWORDS human.

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 714)

AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapb-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: L14M11425 row: m column: 11
High quality sequence stop: 711.

FEATURES
Location/Qualifiers
source 1..714

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5171122"

/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"

/note="Organ: brain; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC Library."

BASE COUNT 152 a 231 c 202 g 129 t
ORIGIN

Query Match 83.8%; Score 482; DB 13; Length 714;
Best Local Similarity 99.8%; Pred. No. 1.1e-108;
Matches 493; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 82 GGATCGCTCTCAAGAGCGTCGCGAGGAGGAGGACCAATCTGTAGCTGGAAGCTGG 141
|||||
Db 81 GGATCGCTCTCAAGAGCGTCGCGAGGAGGAGGACCAATCTGTAGCTGGAAGCTGG 140
|||||

QY 142 GAGCTCCGAAGCGCATCATCTCTGACTTGGTACAGCCACTGTACCCATTTCT 201
|||||
Db 141 GAGCTCCGAAGCGCATCATCTCTGACTTGGTACAGCCACTGTACCCATTTCT 200
|||||

QY 202 GGAGATGTTTCCCGAGTATCCACTGGCTCCCGAGCGCCAGTTTGGAGCAGGGGTG 261
|||||
Db 201 GGAGATGTTTCCCGAGTATCCACTGGCTCCCGAGCGCCAGTTTGGAGCAGGGGTG 260
|||||

QY 262 CCGCAACAGCAGAGTCTCTGACCTGACAGGAGCGCAGTTGGAACCGGGGAGCAG 321
|||||
Db 261 CCGCAACAGCAGAGTCTCTGACCTGACAGGAGCGCAGTTGGAACCGGGGAGCAG 319
|||||

QY 322 AGCCAGGTGGCCAGGTACCAATGGCAATGATGACGGGGGTCTATGACTATCACT 381
|||||
Db 320 AGCCAGGTGGCCAGGTACCAATGGCAATGATGACGGGGGTCTATGACTATCACT 379
|||||

QY 382 GGCACATCTACATCTACATGGACGACTGAGGGGAGCAGCGGGTCTTGAGACCTC 441
|||||
Db 380 GGCACATCTACATCTACATGGACGACTGAGGGGAGCAGCGGGTCTTGAGACCTC 439
|||||

QY 442 CCAGCTACCCCGAAGCTTCATACCCATCCCGAAGAGGGGACCCCTGGCCCTCCCGGG 501
|||||
Db 440 CCAGCTACCCCGAAGCTTCATACCCATCCCGAAGAGGGGACCCCTGGCCCTCCCGGG 499
|||||

QY 502 CTCTCTACACCCACAGGAAGATGGCAAGCTTGGCAGCTAGCGGAGACAGCACTGT 561
|||||
Db 500 CTCTCTACACCCACAGGAAGATGGCAAGCTTGGCAGCTAGCGGAGACAGCACTGT 559
|||||

QY 562 GGTGCCACACCTC 575
|||||
Db 560 GGTGCCACACCTC 573
|||||

RESULT 8
BE740143

LOCUS BE740143

DEFINITION 601595085F1 NIH_MGC_9 Homo sapiens cdna clone IMAGE:3949148 5',
mRNA sequence.

ACCESSION BE740143

VERSION BE740143.1 GI:10154135

LOCUS BE740143

DEFINITION 601595085F1 NIH_MGC_9 Homo sapiens cdna clone IMAGE:3949148 5',
mRNA sequence.

ACCESSION BE740143

VERSION BE740143.1 GI:10154135

LOCUS BE740143

DEFINITION 601595085F1 NIH_MGC_9 Homo sapiens cdna clone IMAGE:3949148 5',
mRNA sequence.

ACCESSION BE740143

VERSION BE740143.1 GI:10154135

LOCUS BE740143

DEFINITION 601595085F1 NIH_MGC_9 Homo sapiens cdna clone IMAGE:3949148 5',
mRNA sequence.

ACCESSION BE740143

VERSION BE740143.1 GI:10154135


```
Db 241 CCACGGTACCAATGGCAATTCATGTCACCGCGGGTCTATGACTATCACTGGCAACATCTA 300
|||||
Qy 393 CATCTACAATGACAGTACTGGGGGACACCGGGTCTCTGGAGACCTCCAGCTACCC 452
|||||
Db 301 CATCTACAATGACCAAGTACTGGGGGACACCGGGTCTCTGGAGACCTCCAGCTACCC 360
|||||
Qy 453 CGAACCTCCATACCCCAATTCCTCCGAAGAGGGGACCTCGCCCTCCCGGGCTCTCTACAC 512
|||||
Db 361 CGAACCTCCATACCCCAATTCCTCCGAAGAGGGGACCTCGCCCTCCCGGGCTCTCTACAC 420
|||||
Qy 513 CCACAGGAAGATGGCAAGCTTGGCACCTAGCGAGAGACAGACTGTGGTGCCACAC 572
|||||
Db 421 CCACAGGAAGATGGCAAGCTTGGCACCTAGCGAGAGACAGACTGTGGTGCCACAC 480
|||||
Qy 573 CTC 575
|||
Db 481 CTC 483
|||

RESULT 10
LOCUS BE769498 900 bp mRNA linear EST 25-SEP-2001
DEFINITION 603059047F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5208394 5',
mRNA sequence.
ACCESSION BE769498
VERSION BE769498.1 GI:15761076
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 900)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1522 row: n column: 11
High quality sequence stop: 854.
FEATURES
Location/Qualifiers
1..900
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5208394"
/clone_lib="NIH_MGC_122"
/lab_host="DH10B"
/Note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
Site 1: NotI; Site 2: EcoRV (destroyed); RNA source:
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC Library."
BASE COUNT 199 a 298 c 246 g 157 t
ORIGIN
Query Match 83.5%; Score 480.4; DB 13; Length 900;
Best Local Similarity 99.6%; Pred. No. 3e-108;
Matches 492; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Qy 82 GGATCGCTGCTCAAGAGCGCTCGCAGGAGAGGGACCAATCTCTAGCTGGAAGCTGG 141
|||||
```

```
Db 269 GGATCGCTGCTCAAGAGCGCTCGCAGGAGAGGGACCAATCTCTAGCTGGAAGCTGG 328
|||||
Qy 142 GAGCCTCCGAAGGCCCATCATCTACTTCCCTGACTTGGTACAGCACTGCTACCAATTTCT 201
|||||
Db 329 GAGCCTCCGAAGGCCCATCATCTACTTCCCTGACTTGGTACAGCACTGCTACCAATTTCT 388
|||||
Qy 202 GGAGATGTTTCCCCAGATATCCACTGGGCTCCCGCAGCCCCAGTTTGGAGGCAGGGGTG 261
|||||
Db 389 GGAGATGTTTCCCCAGATATCCACTGGGCTCCCGCAGCCCCAGTTTGGAGGCAGGGGTG 448
|||||
Qy 262 CGCAACACACAGAGTCTCTGGACCTGACAGGAGCCCGAGTTGGAACCCCGGGGAGCAG 321
|||||
Db 449 CGCAACACACAGAGTCTCTGGACCTGACCA-GGAGCCCGAGTTGGAACCCCGGGGAGCAG 507
|||||
Qy 322 AGCAGGTGGCCCGACGGTACCAATGCAATTCATGTCAACGGGGGGTCTATGACTATCACT 381
|||||
Db 508 AGCAGGTGGCCCGACGGTACCAATGCAATTCATGTCAACGGGGGGTCTATGACTATCACT 567
|||||
Qy 382 GGCACATCTACATCTACAATGACACAGTACTGGGGGGACACCCGGGTCTCTGGAGACCTC 441
|||||
Db 568 GGCACATCTACATCTACAATGACACAGTACTGGGGGGACACCCGGGTCTCTGGAGACCTC 627
|||||
Qy 442 CGAGCTACCCCGAAGCTCATACCCATTCCTCGAAGAGGGGACCTTGGCCCTCCCGGG 501
|||||
Db 628 CCAGCTACCCCGAAGCTCATACCCATTCCTCGAAGAGGGGACCTTGGCCCTCCCGGG 687
|||||
Qy 502 CTCTCTACACCCCGACAGGAGATGCAAGGCTTGGACCTAGCGAGAGACAGACACTGT 561
|||||
Db 688 CTCTCTACAGCCCGACAGGAGATGCAAGGCTTGGACCTAGCGAGAGACAGACACTGT 747
|||||
Qy 562 GGTGCGCACACCCCTC 575
|||||
Db 748 GGTGCGCACACCCCTC 761
|||||

RESULT 11
LOCUS BE789430 749 bp mRNA linear EST 20-OCT-2000
DEFINITION 601482267F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3884920 5',
mRNA sequence.
ACCESSION BE789430
VERSION BE789430.1 GI:10210628
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 749)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: DCTD/DRP/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9659 row: e column: 17
High quality sequence stop: 743.
FEATURES
Location/Qualifiers
1..749
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3884920"
/clone_lib="NIH_MGC_68"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.8 kb. Library constructed by Life
```

```

BASE COUNT      167 a      241 c      220 g      121 t
ORIGIN

Query Match      78.7%; Score 452.8; DB 12; Length 749;
Best Local Similarity 99.4%; Pred. No. 1.8e-101;
Matches 465; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 109 GGAGAGGGACCCAA-TCTGTAGCTGGAGCTGGGAGCTCCGAAAGGCCATCCATCTT 167
|
|
|
Db 1 GGAGAGGGACCCAAAGTCTGTAGCTGGAGCTCCGAAAGGCCATCCATCTT 50
|
|
|
QY 168 CCTGACTTGGTACAGCCACTCTACCCATTTCTGGAGATGTTCCCGAGTATCCACTGG 227
|
|
|
Db 61 CCTGACTTGGTACAGCCACTCTACCCATTTCTGGAGATGTTCCCGAGTATCCACTGG 120
|
|
|
QY 228 GCTCCCGGAGCCCGAGTTTGGAGGAGGGGTGCGCAACAGCAGAGTCTCTGGAGCT 287
|
|
|
Db 121 GCTCCCGGAGCCCGAGTTTGGAGGAGGGGTGCGCAACAGCAGAGTCTCTGGAGCT 180
|
|
|
QY 288 GACCAGGAGCCGAGTTGGAAACCCGGGAGCAGAGCTGGCCCGACGGTACCAATGG 347
|
|
|
Db 181 GACCAGGAGCCGAGTTGGAAACCCGGGAGCAGAGCTGGCCCGACGGTACCAATGG 240
|
|
|
QY 348 CATTGATGTCACCGCGGTCTATGACTATCTACTGGCAACATCTACATCTACATGGAC 407
|
|
|
Db 241 CATTGATGTCACCGCGGTCTATGACTATCTACTGGCAACATCTACATCTACATGGAC 300
|
|
|
QY 408 AGTACTGGGGGACACCGGGTCTCTGGAGACCTCCAGCTACCCCGAACCTCCATACCC 467
|
|
|
Db 301 AGTACTGGGGGACACCGGGTCTCTGGAGACCTCCAGCTACCCCGAACCTCCATACCC 360
|
|
|
QY 468 CATTCCGAAAGAGGGGACCTGGCCCTCCCGGGGTCTCTACACCCCGACAGGAATGG 527
|
|
|
Db 361 CATTCCGAAAGAGGGGACCTGGCCCTCCCGGGGTCTCTACACCCCGACAGGAATGG 420
|
|
|
QY 528 CAAGGCTTGGACCTAGCGGAGACAGACACTGTGTGCCACACCTC 575
|
|
|
Db 421 CAAGGCTTGGACCTAGCGGAGACAGACACTGTGTGCCACACCTC 468

RESULT 12
BG331666      735 bp      mRNA      linear      EST 27-FEB-2001
LOCUS      602432673F1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:4550298 5',
DEFINITION      mRNA sequence.
ACCESSION      BG331666
VERSION      BG331666.1 GI:13138104
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF/Gazdar
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCMI243 row: a column: 19
High quality sequence stop: 709.
Location/Qualifiers
1. 735
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4550298"

FEATURES
source

```

```

/clone_lib="NIH_MGC_18"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dr priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC Library."
BASE COUNT      165 a      232 c      219 g      119 t
ORIGIN

Query Match      77.4%; Score 445; DB 12; Length 735;
Best Local Similarity 99.6%; Pred. No. 1.5e-99;
Matches 467; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 108 GGGAGAGGGACCCCAATCTGTAGCTGGAA-GCTGGAGCCCTCCGAAGGCCATCCATCT 166
|
|
|
Db 2 GGGAGAGGGACCCCAATCTGTAGCTGGAAAGCTGGAGCCCTCCGAAGGCCATCCATCT 51
|
|
|
QY 167 TCCCTGACTTGGTACAGCCACTGTACCCATTTCTGGAGATGTTTCCCGAGTATCCACTG 226
|
|
|
Db 62 TCCCTGACTTGGTACAGCCACTGTACCCATTTCTGGAGATGTTTCCCGAGTATCCACTG 121
|
|
|
QY 227 GGCTCCCGGAGCCCGAGTTTGGAGGAGGGGTGCGCAACAGCAGAGTCTCTGGACC 286
|
|
|
Db 122 GGCTCCCGGAGCCCGAGTTTGGAGGAGGGGTGCGCAACAGCAGAGTCTCTGGACC 181
|
|
|
QY 287 TGACAGGAGCCCGAGTTGGAACCCGGGAGCAGAGCCAGGTGGCCCGACGTACCAATG 346
|
|
|
Db 182 TGACCA-GGAGCCCGAGTTGGAACCCGGGAGCAGAGCCAGGTGGCCCGACGTACCAATG 240
|
|
|
QY 347 GCATTGATGTACCGCGGGTCTATGACTATCTACTGGAACATCTACATCTACATGGAC 406
|
|
|
Db 241 GCATTGATGTACCGCGGGTCTATGACTATCTACTGGAACATCTACATCTACATGGAC 300
|
|
|
QY 407 CAGTACTGGGGGACACCGGGTCTCTGGAGACCTCCAGCTACCCCGAACCTCCATACC 466
|
|
|
Db 301 CAGTACTGGGGGACACCGGGTCTCTGGAGACCTCCAGCTACCCCGAACCTCCATACC 360
|
|
|
QY 467 CCATTCCGAAAGAGGGGACCTGGCCCTCCCGGGGTCTCTACACCCCGACAGGAATG 526
|
|
|
Db 361 CCATTCCGAAAGAGGGGACCTGGCCCTCCCGGGGTCTCTACACCCCGACAGGAATG 420
|
|
|
QY 527 GCAAGGCTTGGACCTAGCGGAGACAGACACTGTGTGCCACACCTC 575
|
|
|
Db 421 GCAAGGCTTGGACCTAGCGGAGACAGACACTGTGTGCCACACCTC 469

RESULT 13
BG331666      606 bp      mRNA      linear      EST 16-OCT-2001
LOCUS      603063085F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5212148 5',
DEFINITION      mRNA sequence.
ACCESSION      BG331666
VERSION      BG331666.1 GI:16168543
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov
Plate: LLAM1532 row: j column: 21
High quality sequence stop: 604.

FEATURES

Location/Qualifiers
1. .606
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5212148"
/clone_lib="NIH_MGC_118"
/tissue_type="Leukocyte"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH_MGC Library."
137 a 194 c 173 g 101 t 1 others

BASE COUNT 137 a 194 c 173 g 101 t 1 others
ORIGIN
Query Match 77.4%; Score 444.8; DB 13; Length 606;
Best Local Similarity 99.1%; Pred. No. 1.6e-99;
Matches 457; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 115 GGACCAATCTCTAGCTGGAAGCTGGAGCTCCGAGGCCATCCATCTTCCCTGAC 174
DB 10 GGGACCAATCTCTAGCTGGAAGCTGGAGCTCCGAGGCCATCCATCTTCCCTGAC 69
QY 175 TTGGTACACCACTGTACCCATTTCTGGAGATGTTTCCCACTATCCACTGGGCTCCC 234
DB 70 TTGGTACACCACTGTACCCATTTCTGGAGATGTTTCCCACTATCCACTGGGCTCCC 129
QY 235 GCAGCCCCAGTTTGGAGGAGGGGTGCGCAACACAGACAGTCTCTGGACCTGACGAG 294
DB 130 GCAGCCCCAGTTTGGAGGCA -GGGTGCGCAACACAGACAGTCTCTGGACCTGACGAG 188
QY 295 GAGCCGCACTGGAACCCGGGAGCAGACCCAGGTGGCCACGGTACCAGTATCCAT 354
DB 189 GAGCCGCACTGGAACCCGGGAGCAGACCCAGGTGGCCACGGTACCAGTATCCAT 248
QY 355 GTACCCGCGGGTCTATGACTATCACTGCACATCTACATCTACATGGACGAGTAC 414
DB 249 GTACCCGCGGGTCTATGACTATCACTGCACATCTACATCTACATGGACGAGTAC 308
QY 415 GGGGACCAACCGGGTCTCTGGAGACCTCCAGCTACCCCGGAGCCTCCATACCCCATTC 474
DB 309 GGGGACCAACCGGGTCTCTGGAGACCTCCAGCTACCCCGGAGCCTCCATACCCCATTC 368
QY 475 GAAGAGGGGACCTGGCCCTCCGGGCTCTCTACACCCACAGGAAGTGGCAAGGCT 534
DB 369 GAAGAGGGGACCTGGCCCTCCGGGCTCTCTACACCCACAGGAAGTGGCAAGGCT 428
QY 535 TGSCACCTAGCGGAGACAGACACTCTGGTGGCACACCCCTC 575
DB 429 TGGCCTTACGGGAGACAGACACTCTGGTGGCACACCCCTC 469

RESULT 14
BG576875
LOCUS BG576875 700 bp mRNA linear EST 10-APR-2001
DEFINITION 60259233F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:4707753',
mRNA sequence.
ACCESSION. BG576875
VERSION BG576875.1 GI:13584528
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

1 (bases 1 to 700)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-re@mail.nih.gov
Tissue procurement: DCTD/Drp
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10578 row: b column: 10
High quality sequence stop: 673.

FEATURES

Location/Qualifiers
1. .700
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4707753"
/clone_lib="NIH_MGC_87"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.383 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
160 a 223 c 189 g 127 t 1 others

BASE COUNT 160 a 223 c 189 g 127 t 1 others
ORIGIN
Query Match 76.0%; Score 437.2; DB 12; Length 700;
Best Local Similarity 98.6%; Pred. No. 1.3e-97;
Matches 483; Conservative 0; Mismatches 3; Indels 4; Gaps 4;
QY 82 GGATCGCTGTCAAGAGGGTCCGAGGAGAGGAGCCCAATCCTGTAGCTGGAAGCTGG 141
DB 84 GGATCGCTGTCAAGAGGGTCCGAGGAGAGGAGCCCAATCCTGTAGCTGGAAGCTGG 143
QY 142 GAGCTCCGAAAGGCCCATCACTTCCCTGACTTGGTACAGCCACTGTACCACTTCT 201
DB 144 GAGCTCCGAAAGGCCCATCACTTCCCTGACTTGGTACAGCCACTGTACCACTTCT 203
QY 202 GGAGATGTTTCCCGAGTATCCACTGGGCTCCCGCAGCCCGCAGTTTGGAGGAGGGGTG 261
DB 204 GGAGATGTTTCCCGAGTATCCACTGGGCTCCCGCAGCCCGCAGTTTGGAGGAGGGGTG 263
QY 262 CCSCAACAGCAGAGTCTCTGGACCTGACACAGGAGCCGAGTTGGAACCCGGGAGCAG 321
DB 264 CCSCAACAGCAGAGTCTCTGGACCTGACACAGGAGCCGAGTTGGAACCCGGGAGCAG 323
QY 322 AGCCAGGTGGCCACCGGTACCAATGGCATTCATGTACCGGGGGGTCTATGACTATCACT 381
DB 324 AGCCAGGTGGCCACCGGTACCAATGGCATTCATGTACCGGGGGGTCTATGACTATCACT 383
QY 382 GGCAAC -ATCTACATCTACAATGGACAGTCTGGGGGAGCCACCGGGTCTCTGGAGACCT 440
DB 384 GGCAACATCTACATCTACATGGACAGTCTGGGGGAGCCACCGGGGTCTCTGGAGACCT 443
QY 441 CCAGCTACCCCGAAACCTCCATACCCCATTCGCCGAAGAGGGGAGCCCTGGCCCTCCCGG 500
DB 444 CCAGCTA -CCCGGAACCTCCATACCCCATTCGCCGAAGAGGGGAGCCCTGGCCCTCCCGG 502
QY 501 GCTCTCTACA -CCCCACGAGAGATGGCAAGCTTGGCACTTAGC -GGAGACAGAGCAG 558
DB 503 GCTCTCTACAGCACCACAGGAAGTGGCAAGCTTGGCACTTAGCAGAGAGAGCAG 562
QY 559 TGTGGTGGCA 568
DB 563 TGTGGTGGCA 572

RESULT 15

BC403820
LOCUS
DEFINITION
602419660f1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4526655 5',
mRNA sequence.

ACCESSION
BG403820

VERSION
BG403820.1 GI:13297268

KEYWORDS
EST.

SOURCE
human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 598)

NIH-MGC <http://mhc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1AM10434 row: h column: 16

High quality sequence stop: 596.

FEATURES

source

1..598

Location/Qualifiers

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4526655"

/clone_lib="NIH_MGC_93"

/tissue_type="transitional cell papilloma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: bladder; Vector: pCMV-SPORT6; Site_1: NotI;

Site_2: SalI; Cloned unidirectionally; oligo-dr primed.

Average insert size 1.7 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

BASE COUNT 133 a 193 c 171 g 101 t

ORIGIN

Query Match 75.1%; Score 432; DB 12; Length 598;

Best Local Similarity 99.6%; Pred. No. 2.3e-96;

Matches 454; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

OY 121 AATCCTGTAGCTGAGCTGGAGCTCCGAGGCCATCCAGCCCATCTACTTCCCTGACTTGGTA 180

|||||

Db 1 AATCCTGTAGCTGAGCTGGAGCTCCGAGGCCATCCAGCCCATCTACTTCCCTGACTTGGTA 60

|||||

OY 181 CAGCCACTGCTACCATTTCTGGAGATGTTTCCCGAGTATCCACTGGGCTCCCGCAGCC 240

|||||

Db 61 CAGCCACTGCTACCATTTCTGGAGATGTTTCCCGAGTATCCACTGGGCTCCCGCAGCC 120

|||||

OY 241 CCAGTTTGGAGGAGGGGTCCCGACAGCAGAGTCTCTGGACCTGACAGGAGCGC 300

|||||

Db 121 CCAG--TTTGGAGGAGGGGTCCCGACAGCAGAGTCTCTGGACCTGACAGGAGCGC 179

|||||

OY 301 CAGTTGAACCCGGGAGCAGCAGCGTGGCCCGCAGCTACCAATGGCATTCATCTACC 360

|||||

Db 180 CAGTTGAACCCGGGAGCAGCAGCGTGGCCCGCAGCTACCAATGGCATTCATCTACC 239

|||||

OY 361 GCGGGGTCTATGACTATCTACCTGGCAACATCTACATGACCATGGACCACTACTGGGGGA 420

|||||

Db 240 GCGGGGTCTATGACTATCTACCTGGCAACATCTACATGACCATGGACCACTACTGGGGGA 299

|||||

OY 421 CCACGGGTCTTGAGACCTCCAGCTACCCCGCAACTCCATACCCCATTCGCCGAGAG 480

|||||

Db 300 CCACGGGTCTTGAGACCTCCAGCTACCCCGCAACTCCATACCCCATTCGCCGAGAG 359

|||||

OY 481 GGGGACCTGGCCCTCCCGGCTCTCTACACCCCGAGGAGATGCAAGGCT-TGGCA 539

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

Db 360 GGGACCCCTGGCCCTCCCGGCTCTCTACACCCCGAGGAGATGGCAAGCTCTGGCA 419

OY 540 CCTAGCGGAGACAGAGCACTGTGTGGCCACACCCCTC 575

|||||

Db 420 CCTAGCGGAGACAGAGCACTGTGTGGCCACACCCCTC 455

|||||

Search completed: April 15, 2003, 16:26:30

Job time : 669.156 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nuclease - nucleic search, using sw model

Run on: April 15, 2003, 13:11:20 ; Search time 18.528 seconds
(without alignments)
9517.464 Million cell updates/sec

Title: US-09-917-372-2_COPY_804_1378

Perfect score: 575

Sequence: 1 ctgtgtgaggcagctccagg.....cactgtgtgtgcacacccctc 575

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*

2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*

3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*

5: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq.*

6: /cgn2_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	35.2	6.1	111282	4	US-09-754-250-3
2	34.2	5.9	268	1	US-08-039-137-16
3	34	5.9	3468	1	US-07-951-715A-2
4	34	5.9	3468	2	US-08-459-448A-2
5	34	5.9	3468	3	US-08-459-595A-2
6	34	5.9	3468	3	US-08-459-504B-2
7	34	5.9	3468	3	US-08-459-444-2
8	34	5.9	3468	3	US-09-053-549-3
9	34	5.9	3468	4	US-09-547-422-2
10	33.2	5.8	4086	1	US-08-313-181-1
11	33	5.7	1057	3	US-09-188-930-18
12	33	5.7	1722	4	US-09-143-623-14
13	33	5.7	2004	3	US-09-188-930-230
14	32.8	5.7	376	4	US-09-149-476-201
15	32.4	5.6	2473	2	US-08-794-494-4
16	32.2	5.6	948	1	US-08-254-359A-31
17	32.2	5.6	948	2	US-08-471-056B-31
18	32.2	5.6	948	2	US-08-484-956-31
19	32.2	5.6	948	2	US-08-757-653-31
20	32.2	5.6	948	2	US-08-599-491-31
21	32.2	5.6	948	2	US-08-756-386-31
22	32.2	5.6	948	2	US-08-823-516-26
23	32.2	5.6	948	3	US-08-682-853A-31
24	32.2	5.6	948	3	US-08-759-038-31
25	32.2	5.6	948	3	US-08-758-314-31
26	32.2	5.6	948	4	US-09-350-309-31
27	32.2	5.6	948	4	US-08-520-946-31

28	32.2	5.6	962	1	US-08-073-384C-11	Sequence 11, Appl
29	32.2	5.6	962	1	US-08-254-359A-11	Sequence 11, Appl
30	32.2	5.6	962	1	US-08-483-043-11	Sequence 11, Appl
31	32.2	5.6	962	1	US-08-481-238-11	Sequence 11, Appl
32	32.2	5.6	962	2	US-08-471-066B-11	Sequence 11, Appl
33	32.2	5.6	962	2	US-08-484-956-11	Sequence 11, Appl
34	32.2	5.6	962	2	US-08-757-653-11	Sequence 11, Appl
35	32.2	5.6	962	2	US-08-599-491-11	Sequence 11, Appl
36	32.2	5.6	962	2	US-08-756-386-11	Sequence 11, Appl
37	32.2	5.6	962	2	US-08-823-516-11	Sequence 11, Appl
38	32.2	5.6	962	3	US-08-682-853A-11	Sequence 11, Appl
39	32.2	5.6	962	3	US-08-759-038-11	Sequence 11, Appl
40	32.2	5.6	962	3	US-08-758-314-11	Sequence 11, Appl
41	32.2	5.6	962	4	US-09-350-309-11	Sequence 11, Appl
42	32.2	5.6	962	4	US-08-520-946-11	Sequence 11, Appl
43	32.2	5.6	963	2	US-08-757-653-162	Sequence 162, App
44	32.2	5.6	963	2	US-08-823-516-60	Sequence 60, Appl
45	32.2	5.6	963	3	US-08-759-038-101	Sequence 101, App

ALIGNMENTS

RESULT 1

US-09-754-250-3/c

; Sequence 3, Application US/09754250

; Patent No. 6376225

; GENERAL INFORMATION:

; APPLICANT: WEI, Ming-Hui et al

; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHODIESTERASE

; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN

; TITLE OF INVENTION: PHOSPHODIESTERASE PROTEINS, AND USES THEREOF

; FILE REFERENCE: CL001063

; CURRENT APPLICATION NUMBER: US/09/754,250

; CURRENT FILING DATE: 2001-01-05

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 111282

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(111282)

; OTHER INFORMATION: n = A,T,C or G

US-09-754-250-3

Query Match

Best Local Similarity 6.1%; Score 35.2; DB 4; Length 111282;

Matches 85; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 2 TGGTGGAGCAGCTCCAGGCACTGCCAGTCCGACACCACTGCGAGGAGGAGGCCA 121

Db 29256 TGGGAAAAGAGCCAGGGCGTCCGCCAGTCTCTTGGCTTCAGGGGTGACTTGAGC 61

QY 62 CACTGCCCCAGAGATGTCAGGATCGCTCAAGAGGCGTCGCGAGGAGGAGGCCA 121

Db 29196 ACTCTCTCCAGCCAGCAGTCCGCCAGTCTCAGAGGAGGAGGCCA 121

QY 122 ATCTGTAGCTGGAGCTGGGAGCTCCGAGGCCATCCATCACTATCC 169

Db 29136 CTCTTGCCTCCCTCAAGTTGGAGCAACCTAGAGCCCTGGCCACACTTCC 29089

RESULT 2

US-08-039-137-16

; Sequence 16, Application US/08039137

; Patent No. 5759771

; GENERAL INFORMATION:

; APPLICANT: Tilanus J.G., Marcel

; TITLE OF INVENTION: Method of determining a Genotype by

; TITLE OF INVENTION: Comparing the Nucleotide Sequence of Members of a Gene

; Patent No. 5759771

RESULT 3
US-07-951-715A-2
Sequence 2, Application US/07951715A
Patent No. 5625136
GENERAL INFORMATION:
APPLICANT: Kozziel, Michael G.
APPLICANT: Desai, Nalinil M.
APPLICANT: Lewis, Kelly S.
APPLICANT: Kramer, Vance C.
APPLICANT: Warren, Gregory W.
APPLICANT: Evola, Stephen V.
APPLICANT: Crossland, Lytle D.
APPLICANT: Wright, Martha S.
APPLICANT: Merlin, Ellis J.
APPLICANT: Lamnis, Karen L.

```

Query Match      5.9%; Score 34; DB 1; Length 3468;
Best Local Similarity 51.3%; Pred. No. 2.4;
Matches       7%; Conservative    0; Mismatches   75; Indels     0; Gaps      10;

QY  404  GACCAGTACTGGGGGACACCGGTCTCTGAGAGACTCCACAGCTACGCCCGCAACCTCCAT 463
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  937  GAGTACTACTGAGGGGCCACAGATCATGCCAGCCCCGTGGGCTTCAGCGGCCCCGAG 996

QY  464  ACCCCATTCCGAGAGAGGGGACCCCTGGCCCTCCGGGGCTCTCTACACCCCACACGAGGA 523
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  997  TTCACCTTCCCCCTGTAGGGACCATGGGCACGCGCGCCCCCAGCAGCGGCATCGTGCC 1056

QY  524  ATGCCAAGGCTTGGCACCTAGCGGAGACAGAGA 557
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  1057 CAGCTGGGCCAGGGCGTGTACCGCACCCCTGAGCA 1090

RESULT 4
US-08-459-448A-2
; Sequence 2, Application US/08459448A
; Patent No. 5859336
```



```

CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Pace, Gary M.
REGISTRATION NUMBER: 40,403
REFERENCE/DOCKET NUMBER: CGC 1995
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8582
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3468 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Synthetic DNA"
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1..3468
OTHER INFORMATION: /product= "Full-length pure maize
OTHER INFORMATION: optimized synthetic Bt"
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3465
US-09-053-549-3

Query Match 5.9%; Score 34; DB 3; Length 3468;
Best Local Similarity 51.3%; Pred. No. 2.4;
Matches 79; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 404 GACCAGTACTGGGGGACCACCGGGTCCTGGAGACCTCCAGCTACCCCGGAACCTCCAT 463
Db 937 GAGTACTAGTGGAGCGGCCACAGATCATGCCAGCCCCCGTGGCTTCACGGCGCCGGAG 996
QY 464 ACCCATTTCCCGAAGAGGGGACCCCTGGCCCTCCCGGGCTCTCTACACCCACCAGGAAG 523
Db 997 TTCACCTTCCCTTGTAAGCGCACCATGGCAACGGCCCCCCCCCAGCAGCGCATCGTGGCC 1056
QY 524 ATGGCAAGGCTTGGCACCTAGCGAGACAGCA 557
Db 1057 CAGCTGGGCGCAGGGCGGTGTACCGACCCCTGAGCA 1090

RESULT 9
US-09-547-422-2
; Sequence 2, Application US/09547422
; Patent No. 6320100
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; Desai, Nalini M.
; Lewis, Kelly S.
; Krammer, Vance C.
; Warren, Gregory W.
; Evola, Stephen V.
; Crossland, Lyle D.
; Wright, Martha S.
; Merlin, Ellis J.
; Launis, Karen L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6320100artis Agribusiness Biotechnology Research, Inc
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/547,422
; FILING DATE: 11-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/459,595
; FILING DATE: 02-JUN-1995
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: S-18805H
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8587
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3468 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..3468
; OTHER INFORMATION: /product= "Full-length pure maize
; optimized synthetic Bt"
; /note= "Disclosed in Figure 3 as synlt.mze"
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
; US-09-547-422-2

Query Match 5.9%; Score 34; DB 4; Length 3468;
Best Local Similarity 51.3%; Pred. No. 2.4; Mismatches 0; Indels 0; Gaps 0;
Matches 79; Conservative 0;

QY 404 GACCAGTACTGGGGGACCACCGGGTCTCTGGAGACCTCCAGCTACCCCGGAACCTCCAT 463
Db 937 GAGTACTACTGGAGCGCCACCATGATGCGCAGCCCGTGGCTTCAGCGGCCCGGAG 996
QY 464 ACCCATTCGGAAGAGGGGACCTGGCCCTCCCGGGTCTCTACACCCACCAAG 523
Db 997 TTCACCTTCGCCCTGTACGGCACCATGGCAAGCGCCGCCCGCCAGCGCATCGTGGCC 1056
QY 524 ATGCAAGGCTGTCACCTAGCGGACAGCA 557
Db 1057 CAGTGGCGGCGGCGGTGTACCGACCTGAGCA 1090

RESULT 10
US-08-313-181-1
; Sequence 1, Application US/08313181
; Patent No. 5681735
; GENERAL INFORMATION:
; APPLICANT: Emerson, Charles P.
; APPLICANT: Goldhamer, David J.
; TITLE OF INVENTION: Transcription Control Element for
; FILE OF INVENTION: Increasing Gene Expression in Myoblasts
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman
; STREET: 1601 Market Street, Suite 720
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
;
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/547,422
; FILING DATE: 11-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/459,595
; FILING DATE: 02-JUN-1995
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: S-18805H
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8587
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3468 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..3468
; OTHER INFORMATION: /product= "Full-length pure maize
; optimized synthetic Bt"
; /note= "Disclosed in Figure 3 as synlt.mze"
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
; US-09-547-422-2
```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,181
; FILING DATE: 07-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Reed, Janet E.
; REGISTRATION NUMBER: 36,252
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 563-4100
; TELEFAX: (215) 563-4044
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4086 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-08-313-181-1

Query Match 5.8%; Score 33.2; DB 1; Length 4086;
Best Local Similarity 45.0%; Pred. No. 4.3;
Matches 125; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

QY 228 GCTCCCGGAGCCCCAGTTTGGAGGAGGGTGGCGGACAGACGAGTCTCTGACCT 287
Db 1724 GCGCAACGCGCATCCGCTATATCGAGGCGCTGCTGCGGACGAGCGCGG 1783
QY 288 GACCAGGAGCGCAGTGTGAACCCGCGGAGCAGAGCGAGTGGCCGACCAATGG 347
Db 1784 CCCCTGGCGCCGACGCGGCGCTTATGCGCGGCGCGCTGCCCGGCGCGCGG 1843
QY 348 CATTCATGTACCGGGGGTCTATGACTATCTACTGCAACATCTACATCTACATGACC 407
Db 1844 CGAGCACTACGCGCGGACTCCGACGCGTCCAGCGCGCTCCCACTGTCTCCGACGCAT 1903
QY 408 AGTACTGGGGGACACCGGGTCTTGAGACCTCCAGCTACCCCGAACCTCCATACC 467
Db 1904 GGTAAAGCGCGGACCGACCGAGGAGTGTAGGGCGGCGCTCGGGATATCAGGGAGC 1963
QY 468 CATTCGCGAGAGGGGAGCGCCCTCCCGGCTCT 505
Db 1964 CGTTCCGAGGGGAGGAGCTGSCCTTGGGGAGGTTT 2001

RESULT 11
US-09-188-930-18
; Sequence 18, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Orust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: SeqSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 1057
; TYPE: DNA
; ORGANISM: Rat
; FEATURE:
```

```
; NAME/KEY: unsure
; LOCATION: (45)...(45)
; NAME/KEY: unsure
; LOCATION: (53)...(53)
; NAME/KEY: unsure
; LOCATION: (116)...(116)
; NAME/KEY: unsure
; LOCATION: (118)...(118)
US-09-188-930-18

Query Match
Best Local Similarity 5.7%; Score 33; DB 3; Length 1057;
Matches 78; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 102 TCGCAGGAGGAGGACCAATCTAGCTGGAAGCTGGAGCCTCCGAGGCCCATCC 161
Db 353 TCGCAGCGGTGGATTCGGATCTCAGCGGAAGAGTGTCCTCATTTACAGGCTCTCC 412
QY 162 ATACTTCCCTGACTTGGTACAGCCACTGCTACCCATTTCTGGAGATGTTTCCCCAGTATC 221
Db 413 CTAAGCCCTGGGACAGTTGATCTCCAAGTACTCTCTCGGGGAGTACACCTGTCAATCA 472
QY 222 CACTGGGCTCCCGGAGCCCGCCAGTTTGGAGGC 254
Db 473 CGCAGGCTTTTGGAGGACCCGATACCTGGGGGC 505

RESULT 12
US-09-142-623-14/C
; Sequence 14, Application US/09142623
; Patent No. 6337201
; GENERAL INFORMATION:
; APPLICANT: KOJI YANAI et al.
; TITLE OF INVENTION: -FRUCTOFURANOSIDASE AND ITS GENE, METHOD OF
; TITLE OF INVENTION: ISOLATING -FRUCTOFURANOSIDASE GENE, SYSTEM FOR PRODUCING
; TITLE OF INVENTION: -FRUCTOFURANOSIDASE, AND -FRUCTOFURANOSIDASE VARIANT
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESS: Wenderoth, Lind & Ponack, L.L.P.
; STREET: 2033 K Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/142,623
; FILING DATE: September 10, 1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee Cheng
; REGISTRATION NUMBER: 40,949
; REFERENCE/DOCKET NUMBER: 98-0989*/LC(WMC)/144
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-721-8200
; TELEFAX: 202-721-8250
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1722 base pairs
; TYPE: Nucleic acid
; STRANDEDNESS: Double stranded
; TOPOLOGY: Linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Microorganism: Scopulariopsis brevicaulis IFO4843
; FEATURE:
; NAME/KEY: mat peptide
```

```
; LOCATION: 1...1722
; IDENTIFICATION METHOD: E
US-09-142-623-14

Query Match
Best Local Similarity 5.7%; Score 33; DB 4; Length 1722;
Matches 69; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 345 TGGCATTATGTCACCGGGGCTCTATGACTATCATCTGCGAACATCTACATCTACATGG 404
Db 838 TGGCAAGATTTTACCGTGGGGTTGTAGCCATCTGTCGAACGACAAATTTGTATGA 779
QY 405 ACCAGTACTGGGGGGACCCGCGTCTGTGAGACCTCCAGCTACCCCGAACCTCCATA 464
Db 778 CCTCGAAGTTGTATGCCCGCCAGCCAGTCACCGCTGCCCGAGTTCGAGTTCCTT 719
QY 465 CCCCATTC 473
Db 718 CCTCGTTC 710

RESULT 13
US-09-188-930-230
; Sequence 230, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.101cl
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 230
; LENGTH: 2004
; TYPE: DNA
; ORGANISM: Rat
US-09-188-930-230

Query Match
Best Local Similarity 5.7%; Score 33; DB 3; Length 2004;
Matches 78; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 102 TCGCAGGAGGAGGACCAATCTGTAGCTGGAAGCTGGAGCCTCCGAGGCCCATCC 161
Db 287 TCGCAGCGGTGGATTCGGATTCGATCTGCAGCGGGAAGAGTGTCCCATTTACAGGCTCTCC 346
QY 162 ATACTTCCCTGACTTGGTACAGCCACTGCTACCCATTTCTGGAGATGTTTCCCGAGTATC 221
Db 347 CTAAGCCCTGGGACAGTTGATCTCCAAGTACTCTCTCGGGGAGTACACCTGTCAATCA 406
QY 222 CACTGGGCTCCCGCAGCCCGCCAGTTTGGAGGC 254
Db 407 CGCAGGCTTTTGGAGGACCCGATACCTGGGGGC 439

RESULT 14
US-09-149-476-201/c
; Sequence 201, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
```

EARLIER	FILING DATE:	1997-04-11
EARLIER	APPLICATION NUMBER:	60/043,312
EARLIER	FILING DATE:	1997-04-11
EARLIER	APPLICATION NUMBER:	60/043,313
EARLIER	FILING DATE:	1997-04-11
EARLIER	APPLICATION NUMBER:	60/043,672
EARLIER	FILING DATE:	1997-04-11
EARLIER	APPLICATION NUMBER:	60/043,315
EARLIER	FILING DATE:	1997-04-11
EARLIER	APPLICATION NUMBER:	60/048,974
EARLIER	FILING DATE:	1997-06-06
EARLIER	APPLICATION NUMBER:	60/056,886
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/056,877
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/056,889
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/056,893
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/056,633
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/056,878
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/056,882
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/056,637
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/056,903
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/056,888
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/056,879
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/056,880
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/056,894
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/056,911
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/056,636
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/056,874
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/056,910
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/056,864
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/056,631
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/056,845
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/056,892
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/057,761
EARLIER	FILING DATE:	1997-08-22
EARLIER	APPLICATION NUMBER:	60/047,595
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,599
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,588
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,585
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,586
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,590
EARLIER	FILING DATE:	1997-05-23
EARLIER	APPLICATION NUMBER:	60/047,594
EARLIER	FILING DATE:	1997-05-23

RESULT 15
US-08-794-494-4/c
; Sequence 4, Application US/08794494
; Patent No. 5981833
; GENERAL INFORMATION:
; APPLICANT: Wise, Roger P.
; APPLICANT: Schnable, Patrick S.

Search completed: April 15, 2003, 16:29:52
Job time : 95.528 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 15:24:11 ; Search time 40.0647 Seconds
(without alignments)
12588.918 Million cell updates/sec

Title: US-09-917-372-2_COPY_804_1378

Perfect score: 575
Sequence: 1 ctgtgtgagcagctccagg.....cactgtgtgtccacacctc 575

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 593429 seqs, 438583890 residues

Total number of hits satisfying chosen parameters: 1186858

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	575	100.0	1982	10	US-09-907-372-2
2	561.2	97.6	574	10	US-09-907-372-8
3	494	85.9	2136	9	US-09-954-531-1348, Ap
4	494	85.9	2136	10	US-09-962-436-262
5	494	85.9	2136	10	US-09-880-107-2135
c 6	452	78.6	651	10	US-09-907-372-7
c 7	425	73.9	425	10	US-09-907-372-9
c 8	253.8	44.1	281	10	US-09-867-701-5601
c 9	251.2	43.7	289	10	US-09-867-701-5666
10	240.4	41.8	371	10	US-09-907-372-16
11	198.4	34.5	219	10	US-09-907-372-10
12	172	29.9	279	10	US-09-907-372-11
13	144.2	25.1	206	10	US-09-907-372-13
c 14	101.8	17.7	862	10	US-09-907-372-12
c 15	61.2	10.6	481	10	US-09-867-701-4881
16	39.2	6.8	548	10	US-09-907-372-14
c 17	36.4	6.3	2192	10	US-09-833-381-906
18	36	6.3	471	10	US-09-907-372-15
19	35.4	6.2	809	10	US-09-844-864-22

c 20	35.2	6.1	111282	12	US-10-094-989-3	Sequence 3, Appli
c 21	34.6	6.0	2930	10	US-09-745-763-198	Sequence 198, App
c 22	34.4	6.0	2768	10	US-09-905-983-4	Sequence 4, Appli
c 23	34.4	6.0	2768	10	US-09-905-983-6	Sequence 6, Appli
c 24	34	5.9	3468	9	US-09-988-452-2	Sequence 2, Appli
c 25	34	5.9	6788	9	US-09-870-759-101	Sequence 101, App
c 26	33.6	5.8	33675	10	US-09-921-992-2	Sequence 2, Appli
c 27	33.2	5.8	2855	9	US-10-125-540-594	Sequence 594, App
c 28	33.2	5.8	2855	10	US-09-764-870-594	Sequence 594, App
c 29	33.2	5.8	15061	9	US-10-092-154-991	Sequence 991, App
c 30	33.2	5.8	15061	10	US-09-764-847-991	Sequence 991, App
c 31	33	5.7	1057	9	US-10-152-661-18	Sequence 18, Appli
c 32	33	5.7	1057	9	US-09-866-050A-18	Sequence 18, Appli
c 33	33	5.7	1722	9	US-09-990-385-14	Sequence 14, Appli
c 34	33	5.7	1890	9	US-10-152-661-447	Sequence 447, App
c 35	33	5.7	1890	9	US-09-866-050A-447	Sequence 447, App
c 36	33	5.7	2004	9	US-10-152-661-230	Sequence 230, App
c 37	33	5.7	2004	9	US-09-866-050A-230	Sequence 230, App
c 38	32.8	5.7	376	9	US-09-809-391-201	Sequence 201, App
c 39	32.2	5.6	595	10	US-09-878-574-4557	Sequence 4557, Ap
c 40	32.2	5.6	785	9	US-09-902-941-1868	Sequence 1868, Ap
c 41	32.2	5.6	785	9	US-09-849-626-1868	Sequence 1868, Ap
c 42	32.2	5.6	785	9	US-10-017-754-1868	Sequence 1868, Ap
c 43	32.2	5.6	948	9	US-10-033-297-26	Sequence 26, Appli
c 44	32.2	5.6	948	9	US-10-081-806-31	Sequence 31, Appli
c 45	32.2	5.6	948	9	US-10-074-328-31	Sequence 31, Appli

ALIGNMENTS

RESULT 1

US-09-907-372-2
; Sequence 2, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Warren, Bridget A.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 1982
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 7497867CB1
US-09-907-372-2

Query Match	100.0%	Score	575;	DB	10;	Length	1982;
Best Local Similarity	100.0%	Pred. No.	2.9e-167;				
Matches	575;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	CTGGTGGAGGAGCTCCAGGACCTGCCAGTCCGACACAACTGCAAAATCCATTAGAG	60				
Db	804	CTGGTGGAGGAGCTCCAGGACCTGCCAGTCCGACACAACTGCAAAATCCATTAGAG	863				
Qy	61	CCACTGCCCCAGAGATGTTCAGGATCGTCTCAAGAGCGCTCCGACGAGGAGGAGGCC	120				
Db	864	CCACTGCCCCAGAGATGTTCAGGATCGTCTCAAGAGCGCTCCGACGAGGAGGAGGCC	923				
Qy	121	AATCTGTAGCTGGAAGCTGGGAGCCTCCGAAGGCCCATCTCCCTGACTTGGTA	180				
Db	924	AATCTGTAGCTGGAAGCTGGGAGCCTCCGAAGGCCCATCTCCCTGACTTGGTA	983				
Qy	181	CAGCCACTGCTACCCATTTCTGGAGATGTTTCCCCAGTAGTCCACTGGGCTCCCCGAGCC	240				
Db	984	CAGCCACTGCTACCCATTTCTGGAGATGTTTCCCCAGTAGTCCACTGGGCTCCCCGAGCC	1043				

Db	301	CAGTTGGAAACCCGGGGAGACGACGAGTGGCGCCACGGTACCAATGGCAATTCATGTCAACC	360
QY	361	GGCGGGTCTATGACTATACCTGGCAACATCTACATCTACAATGGACCACTACTGGGGGA	420
Db	361	GGCGGGTCTATGACTATACCTGGCAACATCTACATCTACAATGGACCACTACTGGGGGA	420
QY	421	CCACCGGGTCTCTGGAGACCTCCCAAGCTACCCCGAACCCTCCATCCCAATCCCGGAAGAG	480
Db	421	CCACCGGGTCTCTGGAGACCTCCCAAGCTACCCCGAACCCTCCATCCCAATCCCGGAAGAG	480
QY	481	GGGGACCTTGGCCCTCCCGGGCTCTCTACACCCCAACAGGAAGATGGCAAGCTTGGAC	540
Db	481	GGGGACCTTGGCCCTCCCGGGCTCTCTAAACCCCAACAGGAAGATGGCAAGCTTGGAC	540
QY	541	CTACGGGAGACAGACACTGTGTGTC	566
Db	541	CTACGGGAGACAGACACTGTGTGTC	566

RESULT 3
US-09-954-531-1348
; Sequence 1348, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:

, TITLE OF INVENTION: Gene Sets
 , FILE REFERENCE: 689290-77
 , CURRENT APPLICATION NUMBER: US/09/954,531
 , CURRENT FILING DATE: 2002-03-02
 , PRIOR APPLICATION NUMBER: US/60/233,133
 , PRIOR FILING DATE: 2000-09-18
 , PRIOR APPLICATION NUMBER: US/60/234,009
 , PRIOR FILING DATE: 2000-09-20
 , PRIOR APPLICATION NUMBER: US/60/234,034
 , PRIOR FILING DATE: 2000-09-20
 , PRIOR APPLICATION NUMBER: US/60/234,509
 , PRIOR FILING DATE: 2000-09-22
 , PRIOR APPLICATION NUMBER: US/60/234,567
 , PRIOR FILING DATE: 2000-09-22
 , NUMBER OF SEQ ID NOS: 1392
 , SOFTWARE: PatentIn version 3.0
 , SEQ ID NO 1348
 , LENGTH: 2136
 , TYPE: DNA
 , ORGANISM: Homo sapiens
 , US-09-954-531-1348

Query Match	85.9%	Score	494	DB	9	Length	2136
Best Local Similarity	100.0%	Prod. No.	2.4e-142	Indels	0	Gaps	
Matches	494	Conservative	0	Mismatches	0		
QY	82	GGATCGCTGCTCAACAGCGCTCCGACGGAGAGGGACCCCAATCCTGTAGCTGGGAAGCTGG	1				
Db	943	GGATCGCTGCTCAGAGGCGTCCGACGGAGAGGACCCCAATCCTGTAGCTGGGAAGCTGG	1				
QY	142	GAGCCTCCGAAGGCCCATCCATATCTCCCTGACTTGGTACAGCCACTGCTACCCATTCT	2				
Db	1003	GAGCCTCCGAAGGCCCATCCATATCTCCCTGACTTGGTACAGCCACTGCTACCCATTCT	1				
QY	202	GGAGATGTTTCCCGAGTACCACTGGGCTCCCGACGCCCCAGTTTTGGAGCAGGGGTG	2				
Db	1063	GGAGATGTTTCCCGAGTATCCACTGGGCTCCCGACGCCCCAGTTTTGGAGCAGGGGTG	1				
QY	262	CCGCAACAGCAGAGTCTCTGGACCTGACAGGAGCGCGAGTGGAAACCCGGGAGCAG	3				
Db	1123	CCGCAACAGCAGAGTCTCTGGACCTGACAGGAGCGCGAGTGGAAACCCGGGAGCAG	1				
QY	322	AGCCAGTGGCCACCGGFACCAATGGCATTCATGTACCCGGGGTCTATGACTATCACT	3				
Db	1183	AGCCAGTGGCCACCGGTACCAATGGCATTCATGTACCCGGGGTCTATGACTATCACT	1				
QY	382	GGCACATCTACATCTACAATGGACCACTACTTGGGGGGAACACCGGGTCTCGGAGACCTC	4				



Db 1243 GGCACATATACATGTACATGACAGTACTGGGGGGACACCGGGTCTTGAGACCTC 1302
QY 442 CCAGGTACCCCGAAGCTCCATACCCATTCGCGAGAGGGGACCTGGCCCTCCCGG 501
Db 1303 CCAGGTACCCCGAAGCTCCATACCCATTCGCGAGAGGGGACCTGGCCCTCCCGG 1362
QY 502 CTCTCTACACCCACAGGAGATGCAAGGCTTGGCACCTAGCGGAGACAGACACTGT 561
Db 1363 CTCTCTACACCCACAGGAGATGCAAGGCTTGGCACCTAGCGGAGACAGACACTGT 1422
QY 562 GGTGCCACACCCCTC 575
Db 1423 GGTGCCACACCCCTC 1436

RESULT 4
US-09-962-436-262
; Sequence 262, Application US/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-75
; CURRENT APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,082
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/234,924
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 262
; LENGTH: 2136
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-436-262

Query Match 85.9%; Score 494; DB 10; Length 2136;
Best Local Similarity 100.0%; Pred. No. 2.4e-142; Mismatches 0; Indels 0; Gaps 0;
Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 82 GGATCGCTGCTCAAGAGGGCTCCGAGGAGAGGGAGCCCAATCCTGTAGCTGGAAGCTGG 141
Db 943 GGATCGCTGCTCAAGAGGGCTCCGAGGAGAGGGAGCCCAATCCTGTAGCTGGAAGCTGG 1002
QY 142 GAGCTCCGAGGCGCATCCATACCTTCCCTGACTTGGTACAGCCACTGCTACCCATTCT 201
Db 1003 GAGCTCCGAGGCGCATCCATACCTTCCCTGACTTGGTACAGCCACTGCTACCCATTCT 1062
QY 202 GGAGATGTTTCCCGAGTATCCACTTCCCTGACTTGGTACAGCCACTGCTACCCATTCT 201
Db 1063 GGAGATGTTTCCCGAGTATCCACTTCCCTGACTTGGTACAGCCACTGCTACCCATTCT 1062
QY 262 CGCAACACAGAGTCTCTGACCTGAGCCCTGACAGGAGCCGAGTTGGAACCCCGGGAGCAG 321
Db 1063 GGAGATGTTTCCCGAGTATCCACTTCCCTGACTTGGTACAGCCACTGCTACCCATTCT 1122
QY 262 CGCAACACAGAGTCTCTGACCTGAGCCCTGACAGGAGCCGAGTTGGAACCCCGGGAGCAG 321
Db 1123 CGCAACACAGAGTCTCTGACCTGAGCCCTGACAGGAGCCGAGTTGGAACCCCGGGAGCAG 1182
QY 322 AGCCAGGTGGCCACCGGTACCAATGCAATTCATGTACCGGGGGTCTATGACTATCACT 381
Db 1183 AGCCAGGTGGCCACCGGTACCAATGCAATTCATGTACCGGGGGTCTATGACTATCACT 1242
QY 382 GGCAACATCTACATCTACATGACCACTTGGGGGGACACCGGGTCTTGAGACCTC 441
Db 1243 GGCAACATCTACATCTACATGACCACTTGGGGGGACACCGGGTCTTGAGACCTC 1302
QY 442 CCAGCTACCCCGAAGCTCCATACCCATTCGCGAGAGGGGACCTGGCCCTCCCGG 501
Db 1303 CCAGCTACCCCGAAGCTCCATACCCATTCGCGAGAGGGGACCTGGCCCTCCCGG 1362
QY 502 CTCTCTACACCCACAGGAGATGCAAGGCTTGGCACCTAGCGGAGACAGACACTGT 561
Db 1423 GGTGCCACACCCCTC 1436

Db 1363 CTCTCTACACCCACAGGAGATGCAAGGCTTGGCACCTAGCGGAGACAGACACTGT 1422
QY 562 GGTGCCACACCCCTC 575
Db 1423 GGTGCCACACCCCTC 1436
RESULT 5
US-09-880-107-2135
; Sequence 2135, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2135
; LENGTH: 2136
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 L04270
US-09-880-107-2135

Query Match 85.9%; Score 494; DB 10; Length 2136;
Best Local Similarity 100.0%; Pred. No. 2.4e-142; Mismatches 0; Indels 0; Gaps 0;
Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 82 GGATCGCTGCTCAAGAGGGCTCCGAGGAGAGGGAGCCCAATCCTGTAGCTGGAAGCTGG 141
Db 943 GGATCGCTGCTCAAGAGGGCTCCGAGGAGAGGGAGCCCAATCCTGTAGCTGGAAGCTGG 1002
QY 142 GAGCTCCGAGGCGCATCCATACCTTCCCTGACTTGGTACAGCCACTGCTACCCATTCT 201
Db 1003 GAGCTCCGAGGCGCATCCATACCTTCCCTGACTTGGTACAGCCACTGCTACCCATTCT 1062
QY 202 GGAGATGTTTCCCGAGTATCCACTTCCCTGACTTGGTACAGCCACTGCTACCCATTCT 261
Db 1063 GGAGATGTTTCCCGAGTATCCACTTCCCTGACTTGGTACAGCCACTGCTACCCATTCT 1122
QY 262 CGCAACACAGAGTCTCTGACCTGAGCCCTGACAGGAGCCGAGTTGGAACCCCGGGAGCAG 321
Db 1123 CGCAACACAGAGTCTCTGACCTGAGCCCTGACAGGAGCCGAGTTGGAACCCCGGGAGCAG 1182
QY 322 AGCCAGGTGGCCACCGGTACCAATGCAATTCATGTACCGGGGGTCTATGACTATCACT 381
Db 1183 AGCCAGGTGGCCACCGGTACCAATGCAATTCATGTACCGGGGGTCTATGACTATCACT 1242
QY 382 GGCAACATCTACATCTACATGACCACTTGGGGGGACACCGGGTCTTGAGACCTC 441
Db 1243 GGCAACATCTACATCTACATGACCACTTGGGGGGACACCGGGTCTTGAGACCTC 1302
QY 442 CCAGCTACCCCGAAGCTCCATACCCATTCGCGAGAGGGGACCTGGCCCTCCCGG 501
Db 1303 CCAGCTACCCCGAAGCTCCATACCCATTCGCGAGAGGGGACCTGGCCCTCCCGG 1362
QY 502 CTCTCTACACCCACAGGAGATGCAAGGCTTGGCACCTAGCGGAGACAGACACTGT 561
Db 1363 CTCTCTACACCCACAGGAGATGCAAGGCTTGGCACCTAGCGGAGACAGACACTGT 1422
QY 562 GGTGCCACACCCCTC 575
Db 1423 GGTGCCACACCCCTC 1436

SOFTWARE: PERL Program
SEQ ID NO 9
LENGTH: 425
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. US20020068242A1 7716340H1
US-09-907-372-9

Query Match 73.9%; Score 425; DB 10; Length 425;
Best Local Similarity 100.0%; Pred. No. 2.9e-121;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 ATCCATTAGAGCCACTGCCCCAGAGATGTTCAGGATCGCTGCTCAAGAGGCGTCCGAGG 109
DB 425 ATCCATTAGAGCCACTGCCCCAGAGATGTTCAGGATCGCTGCTCAAGAGGCGTCCGAGG 366
QY 110 GAGAGGACCCCAATCTGTAGCTGGAGCTGGAGCCCTCCGAAGGCCATCCATCTTCC 169
DB 365 GAGAGGACCCCAATCTGTAGCTGGAGCTGGAGCCCTCCGAAGGCCATCCATCTTCC 306
QY 170 CTGACTGTGTACAGCCACTGTCTACCCATTTCTGGAGATGTTTCCCCAGTATCCACTGGGC 229
DB 305 CTGACTGTGTACAGCCACTGTCTACCCATTTCTGGAGATGTTTCCCCAGTATCCACTGGGC 246
QY 230 TCCCGCAGCCCCAGTTTGGAGGAGGGGTGCGCCCAACAGCAGAGTCTCTGGACCTGA 289
DB 245 TCCCGCAGCCCCAGTTTGGAGGAGGGGTGCGCCCAACAGCAGAGTCTCTGGACCTGA 186
QY 290 CCAGGAGCCGAGTTGGAAACCCGGGGAGCAGAGCAGGTGGCCCAACAGTATCCACTGGCA 349
DB 185 CCAGGAGCCGAGTTGGAAACCCGGGGAGCAGAGCAGGTGGCCCAACAGTATCCACTGGCA 126
QY 350 TTCTATCTACCGCGGGTCTATGACTATCTACCTGGCAACATCTACATCTACATGGACCCAG 409
DB 125 TTCTATCTACCGCGGGTCTATGACTATCTACCTGGCAACATCTACATCTACATGGACCCAG 66
QY 410 TACTGGGGGACACCGGGTCTGGAGAGCTCCAGTACCCCGAAGCTCCATACCCCA 469
DB 65 TACTGGGGGACACCGGGTCTGGAGAGCTCCAGTACCCCGAAGCTCCATACCCCA 6

QY 470 TTCCC 474
DB 5 TTCCC 1

RESULT 8
US-09-867-701-5601
Sequence 5601, Application US/09867701
Patent No. US20020132237A1
GENERAL INFORMATION:
APPLICANT: Agiate, Paul A.
APPLICANT: Jones, Robert
APPLICANT: Harlocker, Susan L.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
FILE REFERENCE: 210121.497
CURRENT APPLICATION NUMBER: US/09/867,701
CURRENT FILING DATE: 2001-05-29
NUMBER OF SEQ ID NOS: 10912
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5601
LENGTH: 281
TYPE: DNA
ORGANISM: Homo sapien
US-09-867-701-5601

Query Match 44.1%; Score 253.8; DB 10; Length 281;
Best Local Similarity 97.1%; Pred. No. 1.2e-68;
Matches 269; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 193 CCCATTCTTGGAGATGTTTCCCACTATCCACTGGGTCCCGCAGGCC-CAGTTTGTGA 251

RESULT 6
US-09-907-372-7/c
Sequence 7, Application US/09907372
Patent No. US20020068242A1
GENERAL INFORMATION:
APPLICANT: Lal, Preeti G.
APPLICANT: Warren, Bridget A.
TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
FILE REFERENCE: PC-0050 US
CURRENT APPLICATION NUMBER: US/09/907,372
CURRENT FILING DATE: 2001-07-27
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PERL Program
SEQ ID NO 7
LENGTH: 651
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20020068242A1 7716364H1
US-09-907-372-7

Query Match 78.6%; Score 452; DB 10; Length 651;
Best Local Similarity 99.6%; Pred. No. 1.6e-129;
Matches 474; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1 CTGGTGGAGGAGCTCCAGGCACTGCCAGTCCGACACCACTCCAAAATCCATTAGAG 60
DB 476 CTGGTGGAGGAGCTCCAGGCACTGCCAGTCCGACACCACTCCAAAATCCATTAGAG 417
QY 61 CCAGTCCCGCCAGAGATGTTCAGGATCGCTGCTCAAGAGGCGTCCGAGGAGGAGCC 120
DB 416 CCAGTCCCGCCAGAGATGTTCAGGATCGCTGCTCAAGAGGCGTCCGAGGAGGAGCC 357
QY 121 ATCTGTAGCTGGAGCTGGAGGCTCCGAAGGCCATCCATCTTCCCTGACTTGGA 180
DB 356 ATCTGTAGCTGGAGCTGGAGGCTCCGAAGGCCATCCATCTTCCCTGACTTGGA 297
QY 181 CAGCCACTGTACCCATTTCTGGAGATGTTTCCCAAGTATCCACTGGGTCCCGCAGCC 240
DB 296 CAGCCACTGTACCCATTTCTGGAGATGTTTCCCAAGTATCCACTGGGTCCCGCAGCC 237
QY 241 CAGTTTTGGAGCAGGGTGGCCGACAGAGTCTCTGAGCTGACCAAGGAGCGG 300
DB 236 CAGTTTTGGAGCAGGGTGGCCGACAGAGTCTCTGAGCTGACCAAGGAGCGG 177
QY 301 CAGTTGGAACCCGGGAGCAGAGCAGGTGGCCAGGTACCAATGGCAATTCATGTACC 360
DB 176 CAGTTGGAACCCGGGAGCAGAGCAGGTGGCCAGGTACCAATGGCAATTCATGTACC 117
QY 361 GCGGGTCTATGACTATCTACCTGCAACATCTACATCTCAATGGACCACTACTGGGGA 420
DB 116 GCGGGTCTATGACTATCTACCTGCAACATCTACATCTCAATGGACCACTACTGGGGA 57
QY 421 CCACCGGCTCTGGAGACCTCCC-AGCTACCCCGCAAGCTCC-ATACCCCATTC 474
DB 56 CCACCGGCTCTGGAGACCTCCCAGTACCTACCTCCCGAAGCTCCATACCCCATTC 1

RESULT 7
US-09-907-372-9/c
Sequence 9, Application US/09907372
Patent No. US20020068242A1
GENERAL INFORMATION:
APPLICANT: Lal, Preeti G.
APPLICANT: Warren, Bridget A.
TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
FILE REFERENCE: PC-0050 US
CURRENT APPLICATION NUMBER: US/09/907,372
CURRENT FILING DATE: 2001-07-27
NUMBER OF SEQ ID NOS: 20

QY	374	CTATCACTGGCAACATCTACATCTACAATGGACCACTACTGGGGGAGACACCGGTCCTG	433
Db	9	CGATCACTGSCAACATCTACATCTACAATGGACCACTACTGGGGGGACACCGGTCCTG	68
QY	434	GAGACCTCCAGCTACCCCCCGAACCCTCCATACCCCATCCCGAAGAGGGGGACCCCTGGCC	493
Db	69	GAGACCTTNCAGTACCCCGGAACCTTCCATACCCCATTCCTCCGAGAGGGGGACCCCTGGCC	128
QY	494	CTCCCGGGCTCTCTACACCCCAACCCAGGAAGATGGCAAGCTTGGCACTTAGCGGAGACAG	553
Db	129	CTNCCGGGCTCTCTACACCCCAACCCAGGAAGATGGCAAGCTTGGCACTTAGCGGAGACAG	188
QY	554	AGCACTGTGTGGCCACACCTCT	575
Db	189	AGCACTGTGTGGCCACACCTCT	210

```

RESULT 12
US-09-907-372-11
; Sequence 11, Application US/09907372
; Patent NO. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Warren, Bridget A.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 11
; LENGTH: 279
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 3321983H1
US-09-907-372-11

```

	Query Match	29.9%;	Score 172;	DB 10;	Length 279;
	Best Local Similarity	100.0%;	Pred. No. 1.8e-43;		
	Matches 172;	Conservative	0;	Mismatches 0;	Indels 0;
	Gaps				
QY	404	GACCACTACTGGGGGGACCAACGGGTCTCTGGAGACCTCCCACTACCCGCCCAACCTCCAT	463		
Db	1	GACCACTACTGGGGGGACCAACGGGTCTCTGGAGACCTCCCACTACCCGCCCAACCTCCAT	60		
QY	464	ACCCCAATCCCGAAGAGGGGACCTCGGCCCTCCGGGCTCTCTACACCCCAACCAAGAG	523		
Db	61	ACCCCAATTCGGAAGAGGGGACCTCGGCCCTCCGGGCTCTCTACACCCCAACCAAGAG	120		
QY	524	ATGCGCAAGGCTTGGCACTTAGCGGAGACAGACACTGTGGTGCCACACCCCTC	575		
Db	121	ATGCGCAAGGCTTGGCACTTAGCGGAGACAGACACTGTGGTGCCACACCCCTC	172		

```

RESULT 13
US-09-907-372-13
; Sequence 13, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Warren, Bridget A.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 13
; LENGTH: 206
; TYPE: DNA
; ORGANISM: Rattus norvegicus
;

```

;
;
;
;
US-09-907-37

Query Match 25.1%; Score 144.2; DB 10; Length 206;
Best Local Similarity 81.5%; Pred. No. 6e-35;
Matches 167; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY	280	CTGGACCTGACCAAGGAGCGCGAGTTGGAAACCCGGGGAGCAGAGCCAGTGGCCCCACGGT	339
Db	2	CTGATCCAGGCCCAAGGAGCTGGAGGCTGAGCCCTGGGGAAACATGGCCAGGTGGCCCCACGGT	61
QY	340	ACCAATGGCAATTGATGTCACCCGGCGGGTCTATGACTATACACTGGCAACAATCTACATCTAC	399
Db	62	GCGAATGGCAATTCACCTGACCGGAGGCTCTGTGACTGTCCCGGCNAATATCTACATATAC	121
QY	400	AATGGACCACTACTGGGGGGACCAACCGGGTCTCTGGAGACCTCCAGCTACCCCCGAACCT	459
Db	122	AATGGGCAGTGTGGGGGGAAACACGGGGCCCTGGAGACCTCCAGTCCCTCCCTGAGCCT	181
QY	460	CCATACCCCAATCCCGAAGAGGGGG	484
Db	182	CCATACCCGACTCCCGAAGAGGGAG	206

```

RESULT 14
US-09-907-372-12/c
; Sequence 12, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Warren, Bridget A.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 12
; LENGTH: 862
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 8576918T
; US-09-907-372-12

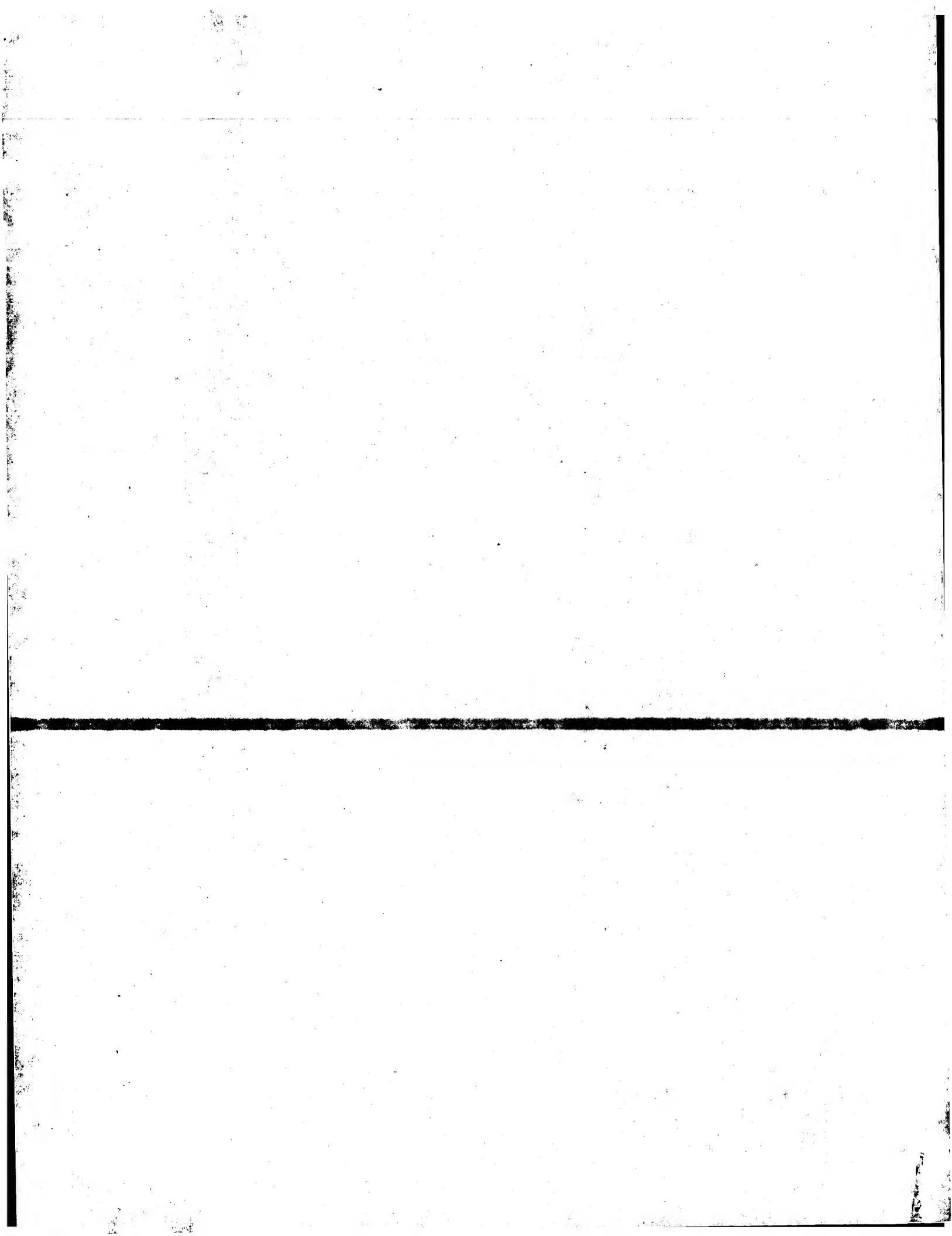
```

[illegible]

RESULT 15
US-09-867-701-4881
; Sequence 4881, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4881
; LENGTH: 481
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-4881

Query Match 10.6%; Score 61.2; DB 10; Length 481;
Best Local Similarity 95.5%; Pred. No. 2.5e-09;
Matches 63; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CTGGTGGAGGCGAGCTCCAGGCGACTGCCGACACACACCTGCAAAATCCATTAGAG 60
|||||
Db 416 CTGGTGGAGGCGAGTCCAGGCGACTGCACAGTCCACACACCTGCAAAATCCATTAGAG 475
|||||
QY 61 CCACTG 66
|||||
Db 476 CCACTG 481

Search completed: April 15, 2003, 18:45:50
Job time : 47.0647 secs



GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2003, 10:39:00 : Search time 5.05967 Seconds
(without alignments)
526.717 Million cell updates/sec

Title: US-09-917-372-1_COPY_216_235
Perfect score: 108
Sequence: 1 PLPPMSGSLKRRPQGECP 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	80	74.1	305	ABP41926	Human ovarian anti
2	52	48.1	570	ABG26356	Novel human diagno
3	51	47.2	127	AAM933585	Human polypeptide,
4	51	47.2	208	ABG12712	Novel human diagno
5	51	47.2	388	ABG70955	S. spinosa protein
6	49	45.4	80	AAU47321	Propionibacterium
7	49	45.4	207	AAU68531	Human novel cytoki
8	49	45.4	601	AAAB43399	Human cancer assoc
9	49	45.4	668	ABB59618	Drosophila melanog
10	49	45.4	891	ABB62720	Drosophila melanog

11	49	45.4	2153	22	AAU33195	Novel human secret
12	48.5	44.9	304	22	ABG07179	Novel human diagno
13	48	44.4	108	21	ABAB1588	Human OREF ORF1352
14	47	43.5	159	21	AGC22742	Zea mays protein f
15	47	43.5	1542	22	ABG71456	Drosophila melanog
16	46.5	43.1	206	22	ABG13470	Novel human diagno
17	46	42.6	77	22	AAW85724	Human immune/haema
18	46	42.6	197	18	AAW23220	Extracellular doma
19	46	42.6	197	20	AAV31326	Human lymphotoxin
20	46	42.6	205	22	ABB17052	Human nervous syst
21	46	42.6	235	22	AAW25470	Human protein sequ
22	46	42.6	371	21	AAW57103	Human prostate can
23	46	42.6	425	22	AAW75454	Human colon cancer
24	46	42.6	471	21	AAW31657	Arabidopsis thalia
25	46	42.6	788	22	AAW39095	Human polypeptide
26	46	42.6	788	23	AAE23718	Human GNK interact
27	46	42.6	820	22	AAW40881	Human polypeptide
28	46	42.6	1064	17	AAW96037	Protein tyrosine k
29	46	42.6	1124	23	AAU96931	Human Jak3 protein
30	46	42.6	1124	23	AAU96933	Human Jak3 protein
31	46	42.6	1124	23	AAU96934	Human Jak3 protein
32	46	42.6	1124	23	AAU96935	Human Jak3 protein
33	46	42.6	1124	23	AAU96936	Human Jak3 protein
34	46	42.6	1124	23	AAU96937	Human Jak3 protein
35	46	42.6	1124	23	AAU96938	Human Jak3 protein
36	46	42.6	1124	23	AAU96939	Human Jak3 protein
37	46	42.6	1124	23	AAU96940	Human Jak3 protein
38	46	42.6	1124	23	AAU96941	Human Jak3 protein
39	46	42.6	1124	23	AAU96942	Human Jak3 protein
40	46	42.6	1124	23	AAU96943	Human Jak3 protein
41	46	42.6	1124	23	AAU96944	Human Jak3 protein
42	46	42.6	1124	23	AAU96945	Human Jak3 protein
43	46	42.6	1124	23	AAU96946	Human Jak3 protein
44	46	42.6	1129	22	ABG20477	Novel human diagno
45	46	42.6	1461	19	AAW64468	Human secreted pro

ALIGNMENTS

RESULT 1
ABP41926
ID ABP41926 standard; Protein; 305 AA.
XX
AC ABP41926;
XX
DT 22-AUG-2002 (first entry)
XX
DE Human ovarian antigen HSABJ44, SEQ ID NO:3058.
XX
KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
KW inflammatory condition; immune disorder; blood disorder;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disorder; urinary system disorder; drug screening;
KW gene therapy; chromosome mapping; forensic analysis;
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW antiinflammatory; gynaecological; reproductive; chromosome 12p13.
OS Homo sapiens.
XX
PN WO200200677-A1.
XX
PD 03-JAN-2002.
XX
PF 07-JUN-2001; 2001WO-US18569.
XX
PR 07-JUN-2000; 2000US-209467P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX

PR 02-MAY-2000; 2000JP-0183765.
XX (HELI-) HELIX RES INST.
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI; 2001-524255/58.
DR N-PSDB; AAK94517.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX
XX Claim 8; SEQ ID NO 3381; 1380pp + sequence listing; English.
XX
CC The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a polypeptide
CC encoded by a full length human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ Sequence 127 AA;

Query Match 47.2%; Score 51; DB 22; Length 127;
Best Local Similarity 50.0%; Pred. No. 7.9;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 PLPPEMSGSLKRRPQEGP 20
||||| :||| :| |
Db 61 PLPPEVRGSLPEGAPWSRAP 80

RESULT 4
ABG12712
ID ABG12712 standard; Protein; 208 AA.
AC
XX ABG12712;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #12703.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX .30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
XX
XX 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR N-PSDB; AAS76899.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess

PT biodiversity -
XX
XX Claim 20; SEQ ID NO 43071; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC the polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 208 AA;

Query Match 47.2%; Score 51; DB 22; Length 208;
Best Local Similarity 55.6%; Pred. No. 13;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 PLPPEMSGSLKRRPQGE 18
||||| :| :|
Db 179 PLPSAMSGSLRPLPEAE 196

RESULT 5
AAB70955
ID AAB70955 standard; Protein; 388 AA.
AC
XX AAB70955;
XX
XX 28-AUG-2001 (first entry)
XX
DE S. spinosa protein fragment encoded by ORF8, SEQ ID 22.
XX
KW Forosamine; trimethylrhamnose; polyketide synthase; biosynthesis;
KW spinosyn; polyketide aglycone; transgenic plant; insect resistance;
KW macrolide; insecticidal; C-C cyclising enzyme.
XX
XX Saccharopolyspora spinosa.
XX
XX DE19957268-A1.
XX
XX 08-MAR-2001.
XX
XX 29-NOV-1999; 99DE-1057268.
XX
XX 27-AUG-1999; 99DE-1040596.
XX
XX (FARB) BAYER AG.
XX
XX Eberz G, Moehrle V, Froede R, Velten R, Salas JA;
XX
XX WPI; 2001-267102/28.
DR N-PSDB; AAF88325.
XX
XX New nucleic acid encoding enzymes for spinosyn biosynthesis, useful for
PT recombinant production of insecticidal spinosyns and their derivatives
PT
XX Claim 39; Page 131-132; 354pp; German.
PS

XX This invention describes a novel method nucleic acid (I) and its encoded
 CC polypeptide (II) containing at least one region that encodes an enzymatic
 CC activity involved in biosynthesis of spinosyns. (I) are used (i) to
 CC identify, inactivate or modulate genes involved in the biosynthesis of
 CC (II); (ii) to generate a library of polypeptide synthases; (iii) for
 CC adding forosamine or trimethylrhamsone to a spinosyn or polypeptide
 CC aglycone; and (iv) for recombinant production of the corresponding
 CC derivatives, including production of (II), their precursors or
 CC and thus have increased resistance to insects. (I) are also useful as
 CC markers for sequencing of the Saccharopolyspora spinosa genome. (II) are
 CC macrolides with insecticidal, but not antibacterial, activity, and can
 CC also be used to raise specific antibodies, useful for identifying
 CC expression clones in a gene bank. Cells transformed with (I) may produce
 CC (II) at significantly increased levels or produce new derivatives of
 CC (II). This sequence represents an S. spinosa C-C cyclising enzyme.
 XX

SQ Sequence 388 AA;

Query Match 47.2%; Score 51; DB 22; Length 388;
 Best Local Similarity 60.0%; Pred. No. 25;
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 PLPPEMSGSLKKRRP 15
 |||||:|:|:|
 Db 27 PLPPGVPTLLRARP 41

RESULT 6

AAU47321
 ID AAU47321 standard; Protein; 80 AA.

AC AAU47321;

DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #8217.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12865.

XX 21-APR-2000; 2000US-199047P.

XX 02-JUN-2000; 2000US-208841P.

XX 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

XX N-PSDB; AAS59538.

XX Propionibacterium acnes polypeptides and nucleic acids useful for

XX vaccinating against and diagnosing infections, especially useful for

XX treating acne vulgaris -

XX Example 1; SEQ ID No 8516; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic

XX polypeptides. The proteins and their associated DNA sequences are used in

XX the treatment, prevention and diagnosis of medical conditions caused by

CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 80 AA;

Query Match 45.4%; Score 49; DB 22; Length 80;
 Best Local Similarity 50.0%; Pred. No. 9.7;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 PLPPEMSGSLKKRRPQGE 18
 | |||:|:|:|
 Db 52 PSPPELAGARLGRHYGGQ 69

RESULT 7

AAU68531

ID AAU68531 standard; Protein; 207 AA.

AC AAU68531;

DT 16-JAN-2002 (first entry)

DE Human novel cytokine encoded by cDNA 790CIP2C_2 #1.

XX Human; cytokine; cell proliferation; cell differentiation;
 KW antiinflammatory; stem cell growth factor; activin; inhibin; cancer;
 KW nervous system disease; neuropathy; Alzheimer's disease;
 KW Parkinson's disease; Huntington's disease; spinal cord disorder;
 KW head trauma; stroke; myeloid cell disorder; lymphoid cell disorder;
 KW platelet disorder; thrombocytopenia; stem cell disorder;
 KW aplastic anaemia; tissue regeneration; wound healing; ulcer;
 KW osteoporosis; osteoarthritis; bone degenerative disorder;
 KW periodontal disease; fibrosis; reperfusion; immune disorder; SCID;
 KW severe combined immunodeficiency; infection; autoimmune disorder;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; allergy;
 KW asthma; coagulation disorder; haemophilia; sepsis; nephritis;
 KW inflammatory bowel disease; food supplement; immunogen.

XX Homo sapiens.

XX WO200175093-A1.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US10484.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX 22-SEP-2000; 2000US-0668680.

XX 23-OCT-2000; 2000US-0695618.

XX 30-NOV-2000; 2000US-0728711.

XX 14-MAR-2001; 2000US-0728711.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J, Xu C;

XX Yang Y, Zaho QA, Chen R, Wang D, Goodrich RW, Liu C, Drmanac RT;

DR WPI; 2001-G26432/72.
 XX N-PSDB; AAS59823.
 XX
 PT New polypeptides and nucleic acids, useful for diagnosis, treatment of
 PT inflammatory, autoimmune, neurological, myeloid or lymphoid cell, bone
 PT degenerative disorders, cancer and promoting wound healing
 XX
 XX
 PS Claim 20; Page 248; 336pp; English.
 XX
 XX The invention relates to isolated human polypeptides (which may be
 CC cytokines) and the polynucleotides encoding them. The protein is useful
 CC for identifying a compound which binds to it (e.g. modulators, agonists
 CC and antagonists). The polynucleotides are useful as an array for mismatch
 CC detection. The proteins and nucleic acids are useful as nutritional
 CC sources or supplements. The protein exhibits activity relating
 CC to cytokine, cell proliferation, cell differentiation, antiinflammatory,
 CC stem cell growth factor activity, immune stimulating or immune
 CC suppressing and activin or inhibin related activities. The proteins (and
 CC antibodies raised against them) and nucleic acids are therefore useful in
 CC the diagnosis and treatment of diseases and disorders such as cancer,
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, spinal cord disorders, head trauma, cerebrovascular
 CC diseases, stroke, myeloid or lymphoid cell disorders, platelet disorders,
 CC thrombocytopaenia, stem cell disorders, aplastic anaemia, for
 CC regeneration of bone, cartilage, tendon, ligament and/or nerve tissue
 CC growth, and in tissue repair, healing of burns, incisions, ulcers, for
 CC treating osteoporosis, osteoarthritis, bone degenerative disorders, or
 CC periodontal disease, lung or liver fibrosis, reperfusion injury in
 CC various tissues, various immune deficiencies and disorders including
 CC severe combined immunodeficiency (SCID), bacterial or fungal infections,
 CC autoimmune disorders (e.g. multiple sclerosis, rheumatoid arthritis,
 CC diabetes mellitus, myasthenia gravis), allergic reactions and conditions,
 CC such as asthma or other respiratory problems, coagulation disorders,
 CC haemophilia, septic shock, sepsis, arthritis, nephritis and inflammatory
 CC bowel disease, viral infection and are useful in altering bodily
 CC characteristics. The present sequence represents a novel protein of the
 CC invention.
 XX
 SQ Sequence 207 AA;
 Query Match 45.4%; Score 49; DB 22; Length 207;
 Best Local Similarity 31.2%; Pred. No. 26;
 Matches 10; Conservative 5; Mismatches 5; Indels 12; Gaps 1;
 QY 1 PLPPEMSGSL-----LKRPPQGEPP 20
 ||||: :| :|:| :|:|
 Db 168 PLPPTAGRAQHSQSPRRALPAKRPQGP 199
 RESULT 8
 AAB43399
 ID AAB43399 standard; Protein; 601 AA.
 XX
 AC AAB43399;
 XX
 XX 08-FEB-2001 (first entry)
 XX
 XX Human cancer associated protein sequence SEQ ID NO:844.
 DE
 XX Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
 KW antidiabetic; antiaesthetic; antirheumatic; antithrombotic; antiviral;
 KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
 KW dermatological; neuroprotective; thrombolytic; coagulant; neurotropic;
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disease; organ rejection;
 KW allergic reaction; graft versus host disorder; autoimmune disorder;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening.
 XX
 OS Homo sapiens.
 XX

PN WO200055350-A1.
 XX
 PD 21-SEP-2000.
 XX
 XX 08-MAR-2000; 2000WO-US05882.
 PF
 XX 12-MAR-1999; 99US-0124270.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Rosen CA, Ruben SM;
 PI
 XX WPI; 2000-587533/55.
 DR N-PSDB; AAC77608.
 XX
 XX Novel isolated nucleic acids comprising sequences encoding peptides
 PT useful for treating or diagnosing e.g. cancer -
 PT
 XX Claim 11; Page 1396-1398; 2352pp; English.
 PS
 CC AAC77607 to AAC78448 encode the human cancer associated proteins given
 CC in AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnery; immunomodulator;
 CC antidiabetic; antiaesthetic; antirheumatic; antithrombotic;
 CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
 CC neurotropic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention.
 XX
 SQ Sequence 601 AA;
 Query Match 45.4%; Score 49; DB 21; Length 601;
 Best Local Similarity 47.4%; Pred. No. 76;
 Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
 QY 2 LPPEMSGSLKRRPQGEPP 20
 ||||: :| :|:| :|:|
 Db 194 LPPPHSSGFLGSRKPEGPP 212
 RESULT 9
 ABB59618
 ID ABB59618 standard; Protein; 668 AA.
 XX
 AC ABB59618;
 XX
 XX 26-MAR-2002 (first entry)
 DT
 XX Drosophila melanogaster polypeptide SEQ ID NO 5646.
 DE
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 KW Drosophila melanogaster.
 OS
 XX WO200171042-A2.
 PN
 XX 27-SEP-2001.
 PD
 XX 23-MAR-2001; 2001WO-US09231.
 PF

CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation; to regulate haematopoiesis; and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and
 CC in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.

XX Sequence 2153 AA;

Query Match 45.4%; Score 49; DB 22; Length 2153;
 Best Local Similarity 47.4%; Pred. No. 2.8e+02;
 Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 LPPEMSGSLKRRPQEGP 20
 ||| | | :|||
 Db 1562 LPPHSSGFLGSKRPEGPG 1580

RESULT 12

ABG07179
 ID ABG07179 standard; Protein; 304 AA.

XX AC ABG07179;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #7170.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX PP WPI; 2001-639362/73.

XX DR N-PSDB; AAS71366.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

XX PS Claim 20; SEQ ID No 37538; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 304 AA;

Query Match 44.9%; Score 48.5; DB 22; Length 304;
 Best Local Similarity 52.0%; Pred. No. 45;
 Matches 13; Conservative 1; Mismatches 4; Indels 7; Gaps 2;

QY 1 PLPPEM-SGSL-RRRPOQE 18
 ||||| | | :|||
 Db 250 PLPPEFSGPALSSCRXGNRRPEGE 274

RESULT 13

AAB41588
 ID AAB41588 standard; Protein; 108 AA.

XX AC AAB41588;

XX DT 08-FEB-2001 (first entry)

XX DE Human ORFX ORF1352 polypeptide sequence SEQ ID NO:2704.

XX KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antinflammatory disease; coagulation;
 KW thrombosis; contraceptive.

XX OS Homo sapiens.

XX PN WO200058473-A2.

XX PD 05-OCT-2000.

XX PF 31-MAR-2000; 2000WO-US08621.

XX PR 31-MAR-1999; 99US-0127607.

XX PR 02-APR-1999; 99US-0127636.

XX PR 05-APR-1999; 99US-0127728.

XX PR 30-MAR-2000; 2000US-0540763.

XX PA (CURA-) CURAGEN CORP.

XX PI Shimkets RA, Leach M;

XX DR WPI; 2000-602362/57.

XX DR N-PSDB; AAC75797.

XX PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -

XX PS Claim 11; Page 1941; 5507pp; English.

XX CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;
 CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;

CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antithyroid; and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance .
 CC coagulation; to inhibit thrombosis; and as a contraceptive.

SQ Sequence 108 AA;

Query Match 44.4%; Score 48; DB 21; Length 108;
 Best Local Similarity 63.2%; Pred. No. 18;
 Matches 12; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

QY 2 LPPEMSGILKRRPOGEP 20
 |||||I-|||I|||
 Db 66 LPPE--GILLRPLLGECP 82

RESULT 14

AAG22742
 ID AAG22742 standard; Protein; 159 AA.

XX AAG22742;

AC AAG22742;

XX 17-OCT-2000 (first entry)

DE Zea mays protein fragment SEQ ID NO: 25788.

KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence; corn.

XX Zea mays subsp. mays.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 28-APR-1999; 99US-0130891.

XX 28-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 30-APR-1999; 99US-0132407.

XX 04-MAY-1999; 99US-0132484.

XX 05-MAY-1999; 99US-0132485.

XX 06-MAY-1999; 99US-0132486.

XX 06-MAY-1999; 99US-0132487.

XX 07-MAY-1999; 99US-0132863.

XX 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 24-JUN-1999; 99US-0140354.

PR 28-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 23-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.

23-JUL-1999; 99US-0145224.
26-JUL-1999; 99US-0145276.
27-JUL-1999; 99US-0145913.
27-JUL-1999; 99US-0145918.
27-JUL-1999; 99US-0145919.
28-JUL-1999; 99US-0145951.
02-AUG-1999; 99US-0146386.
02-AUG-1999; 99US-0146388.
02-AUG-1999; 99US-0146389.
03-AUG-1999; 99US-0147038.
04-AUG-1999; 99US-0147204.
04-AUG-1999; 99US-0147302.
05-AUG-1999; 99US-0147192.
05-AUG-1999; 99US-0147260.
06-AUG-1999; 99US-0147303.
06-AUG-1999; 99US-0147416.
09-AUG-1999; 99US-0147493.
09-AUG-1999; 99US-0147935.
10-AUG-1999; 99US-0147935.
11-AUG-1999; 99US-0148171.
12-AUG-1999; 99US-0148319.
13-AUG-1999; 99US-0148341.
13-AUG-1999; 99US-0148565.
16-AUG-1999; 99US-0148684.
17-AUG-1999; 99US-0149368.
17-AUG-1999; 99US-0149175.
18-AUG-1999; 99US-0149426.
20-AUG-1999; 99US-0149722.
20-AUG-1999; 99US-0149723.
20-AUG-1999; 99US-0149929.
23-AUG-1999; 99US-0149902.
23-AUG-1999; 99US-0149930.
25-AUG-1999; 99US-0150566.
26-AUG-1999; 99US-0150884.
27-AUG-1999; 99US-0151065.
27-AUG-1999; 99US-0151066.
27-AUG-1999; 99US-0151080.
30-AUG-1999; 99US-0151303.
31-AUG-1999; 99US-0151438.
01-SEP-1999; 99US-0151930.
07-SEP-1999; 99US-0152363.
10-SEP-1999; 99US-0153070.
13-SEP-1999; 99US-0153758.
15-SEP-1999; 99US-0154018.
16-SEP-1999; 99US-0154039.
20-SEP-1999; 99US-0154779.
22-SEP-1999; 99US-0155139.
23-SEP-1999; 99US-0155486.
24-SEP-1999; 99US-0155659.
28-SEP-1999; 99US-0156458.
29-SEP-1999; 99US-0156596.
04-OCT-1999; 99US-0157117.
05-OCT-1999; 99US-0157753.
08-OCT-1999; 99US-0157865.
07-OCT-1999; 99US-0158029.
08-OCT-1999; 99US-0158232.
12-OCT-1999; 99US-0158369.
13-OCT-1999; 99US-0159293.
13-OCT-1999; 99US-0159294.
13-OCT-1999; 99US-0159295.
14-OCT-1999; 99US-0159329.
14-OCT-1999; 99US-0159330.
14-OCT-1999; 99US-0159331.
14-OCT-1999; 99US-0159637.
18-OCT-1999; 99US-0159638.
21-OCT-1999; 99US-0160741.
21-OCT-1999; 99US-0160767.
21-OCT-1999; 99US-0160768.
21-OCT-1999; 99US-0160770.
21-OCT-1999; 99US-0160814.
21-OCT-1999; 99US-0160815.
22-OCT-1999; 99US-0160980.
22-OCT-1999; 99US-0160981.

22-OCT-1999; 99US-0160989.
25-OCT-1999; 99US-0161404.
25-OCT-1999; 99US-0161405.
25-OCT-1999; 99US-0161406.
26-OCT-1999; 99US-0161359.
26-OCT-1999; 99US-0161360.
26-OCT-1999; 99US-0161361.
28-OCT-1999; 99US-0161920.
28-OCT-1999; 99US-0161992.
28-OCT-1999; 99US-0161993.
29-OCT-1999; 99US-0162142.

Query Match 43.5%; Score 47; DB 21; Length 159;
Best Local Similarity 47.4%; Pred. No. 38;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 PLPPMSGSLKKRRPQEG 19
| | | | | | | | | |
Db 16 PPPQSSRRSRNPRGTG 34

RESULT 15

ABB71456
ID ABB71456 standard; Protein; 1542 AA.
XX AC ABB71456;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster polypeptide SEQ ID NO 41160.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.
XX OS Drosophila melanogaster.
XX WC200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US09231.
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX DR N-PSDB; ABL15559.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX interactions - genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX PS Disclosure; SEQ ID NO 41160; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABB57737-ABB72072).
XX CC The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 1542 AA;

Query Match 43.5%; Score 47; DB 22; Length 1542;

Best Local Similarity 52.9%; Pred. No. 3.9e+02;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Oy 4 PMSGSLKKRRPQGEF 20
| | : | | | | |
Db 1016 PYMPDHMYPRRPGAGP 1032

Search completed: April 8, 2003, 10:49:07
Job time : 7.05967 secs